



(12)

## EUROPEAN PATENT APPLICATION

(43) Date of publication:  
01.09.1999 Bulletin 1999/35(51) Int. Cl.<sup>6</sup>: G10L 9/14

(21) Application number: 99102792.1

(22) Date of filing: 24.02.1999

(84) Designated Contracting States:  
 AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU  
 MC NL PT SE  
 Designated Extension States:  
 AL LT LV MK RO SI

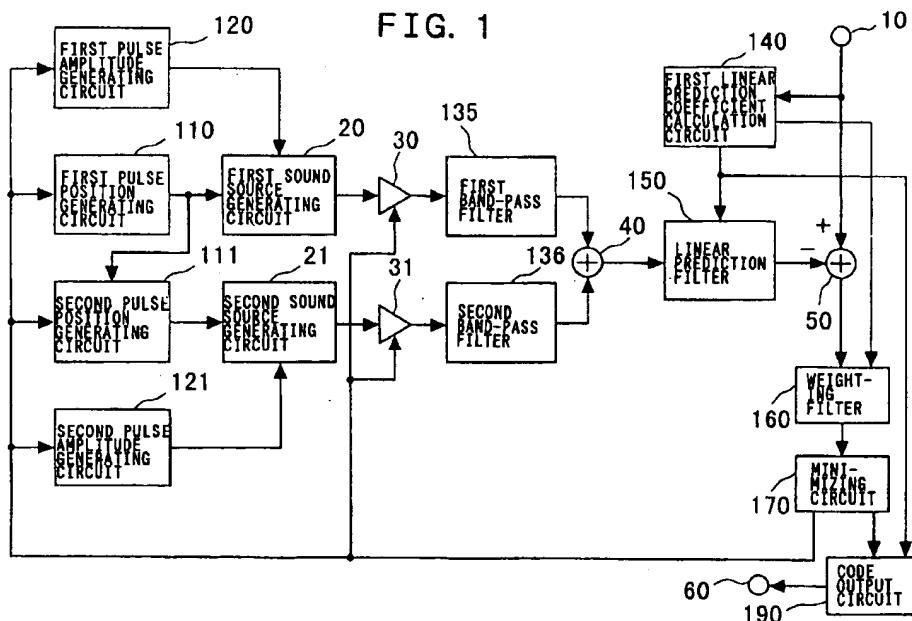
(30) Priority: 27.02.1998 JP 6472198

(71) Applicant: NEC CORPORATION  
Tokyo (JP)(72) Inventor: Murashima, Atsushi  
Minato-ku, Tokyo (JP)(74) Representative:  
 VOSSIUS & PARTNER  
 Siebertstrasse 4  
 81675 München (DE)

## (54) Apparatus for encoding and apparatus for decoding speech and musical signals

(57) A speech and musical signal codec employing a band splitting technique encodes sound source signals of each of a plurality of bands using a small number of bits. The codec includes a second pulse position generating circuit, to which an index output by a minimizing circuit and a first pulse position vector  $P^* = (P_1, P_2, \dots, P_M)$  are input, for revising the first pulse position vector

using a pulse position revision quantity  $d^* = (d_{i1}, d_{i2}, \dots, d_{iM})$  specified by the index and outputting the revised vector to a second sound source generating circuit as a second pulse position vector  $P^* = (P_1+d_{i1}, P_2+d_{i2}, \dots, P_M+d_{iM})$ .



## Description

[0001] This invention relates to an apparatus for encoding and an apparatus for decoding speech and musical signals. More particularly, the invention relates to a coding apparatus and a decoding apparatus for transmitting speech and musical signals at a low bit rate.

[0002] A method of encoding a speech signal by separating the speech signal into a linear prediction filter and its driving sound source signal is used widely as a method of encoding a speech signal efficiently and medium to low bit rates.

[0003] One such method that is typical is CELP (Code-Excited Linear Prediction). With CELP, a linear prediction filter for which linear prediction coefficients obtained by subjecting input speech to linear prediction analysis have been decided is driven by a sound source signal represented by the sum of a signal that represents the speech pitch period and a noise signal, whereby there is obtained a synthesized speech signal (i.e., a reconstructed signal). For a discussion of CELP, see the paper (referred to as "Reference 1") "Code excited linear prediction: High quality speech at very low bit rates" by M. Schroeder et. al (Proc. ICASSP, pp. 937 - 940, 1985).

[0004] A method using a higher-order linear prediction filter representing the complicated spectrum of music is known as a method of improving music encoding performance by CELP. According to this method, the coefficients of a higher-order linear prediction filter are found by applying linear prediction analysis at a high order of from 50 to 100 to a signal obtained by inverse filtering a past reconstructed signal using a linear prediction filter. A signal obtained by inputting a musical signal to the higher-order linear prediction filter is applied to a linear prediction filter to obtain the reconstructed signal.

[0005] As an example of an apparatus for encoding speech and musical signals using a higher-order prediction linear filter, see the paper (referred to as "Reference 2") "Improving the Quality of Musical Signals in CELP Coding", by Sasaki et al. (Acoustical Society of Japan, Spring, 1996 Meeting for Reading Research Papers, Collected Papers, pp. 263 - 264, 1996) and the paper (referred to as "Reference 3") "A 16 Kbit/s Wideband CELP Coder with a High-Order Backward Predictor and its Fast Coefficient Calculation" by M Serizawa et al. (IEEE Workshop on Speech Coding for Telecommunications, pp. 107 - 108, 1997).

[0006] A known method of encoding a sound source signal in CELP involves expressing a sound source signal efficiently by a multipulse signal comprising a plurality of pulses and defined by the positions of the pulses and pulse amplitudes.

[0007] For a discussion of encoding of a sound source signal using a multipulse signal, see the paper (referred to as "Reference 4") "MP-CELP Speech Coding based upon a Multipulse Spectrum Quantized Sound Source

and High-Speed Searching" by Ozawa et. al (Collected Papers A of the Society of Electronic Information Communications, pp. 1655 - 1663, 1996). Further, by adopting a band splitting arrangement using a sound source signal found for each band and a higher-order backward linear prediction filter in an apparatus for encoding speech and musical signals based upon CELP, the ability to encode music is improved.

[0008] With regard to CELP using band splitting, see the paper (referred to as "Reference 5") "Multi-band CELP Coding of Speech and Music" by A. Ubale et al. (IEEE Workshop on Speech Coding for Telecommunications, pp. 101 - 102, 1997).

[0009] Fig. 10 is a block diagram showing an example of the construction of an apparatus for encoding speech and music according to the prior art. For the sake of simplicity, it is assumed here that the number of bands is two.

[0010] As shown in Fig. 10, an input signal (input vector) enters from an input terminal 10. The input signal is generated by sampling a speech or musical signal and gathering a plurality of the samples into a single vector as one frame.

[0011] A first linear prediction coefficient calculation circuit 140 receives the input vector as an input from the input terminal 10. This circuit subjects the input vector to linear prediction analysis, obtains a linear prediction coefficient and quantizes the coefficient. The first linear prediction coefficient calculation circuit 140 outputs the linear prediction coefficient to a weighting filter 160 and outputs an index, which corresponds to a quantized value of the linear prediction coefficient, to a linear prediction filter 150 and to a code output circuit 690.

[0012] A known method of quantizing a linear prediction coefficient involves converting the coefficient to a line spectrum pair (referred to as an "LSP") to effect quantization. For a discussion of the conversion of a linear prediction coefficient to an LSP, see the paper (referred to as "Reference 6") "Speech Information Compression by Line Spectrum Pair (LSP) Speech Analysis Synthesis" by Sugamura et al. (Collected Papers A of the Society of Electronic Information Communications, Vol. J64-A, No. 8, pp. 599 - 606, 1981). In regard to quantization of an LSP, see the paper (referred to as "Reference 7") "Vector Quantization of LSP Parameter Using Running-Mean Interframe Prediction" by Omuro et al. (Collected Papers A of the Society of Electronic Information Communications, Vol. J77-A, No. 3, pp. 303 - 312, 1994).

[0013] A first pulse position generating circuit 610 receives as an input an index that is output by a minimizing circuit 670, generates a first pulse position vector using the position of each pulse specified by the index and outputs this vector to a first sound source generating circuit 20.

[0014] Let M represent the number of pulses and let P1, P2, ..., PM represent the positions of the pulses. The vector P, therefore, is written as follows:

$$= (P_1, P_2, \dots, P_M)$$

(It should be noted that the bar over P indicates that P is a vector.)

[0015] A first pulse amplitude generating circuit 120 has a table in which M-dimensional vectors  $A_j$ ,  $j = 1, \dots, N_A$  have been stored, where  $N_A$  represents the size of the table. The index output by the minimizing circuit 670 enters the first pulse amplitude generating circuit 120, which proceeds to read an M-dimensional vector  $A_j$  corresponding to this index out of the above-mentioned table and outputs this vector to the first sound source generating circuit 20 as a first pulse amplitude vector.

[0016] Letting  $A_{i1}, A_{i2}, \dots, A_{iM}$  represent the amplitude values of the pulses, we have

$$A_j = (A_{i1}, A_{i2}, \dots, A_{iM})$$

[0017] A second pulse position generating circuit 611 receives as an input the index that is output by the minimizing circuit 670, generates a second pulse position vector using the position of each pulse specified by the index and outputs this vector to a second sound source generating circuit 21.

[0018] A second pulse amplitude generating circuit 121 has a table in which M-dimensional vectors  $B_j$ ,  $j = 1, \dots, N_B$  have been stored, where  $N_B$  represents the size of the table.

[0019] The index output by the minimizing circuit 670 enters the second pulse amplitude generating circuit 121, which proceeds to read an M-dimensional vector  $B_j$  corresponding to this index out of the above-mentioned table and outputs this vector to the second sound source generating circuit 21 as a second pulse amplitude vector.

[0020] The first pulse position vector  $P = (P_1, P_2, \dots, P_M)$  output by the first pulse position generating circuit 610 and the first pulse amplitude vector  $A_j = (A_{i1}, A_{i2}, \dots, A_{iM})$  output by the first pulse amplitude generating circuit 120 enter the first sound source generating circuit 20. The first sound source generating circuit 20 outputs an N-dimensional vector for which the values of the  $P_1$ st,  $P_2$ nd, ...,  $P_M$ th elements are  $A_{i1}, A_{i2}, \dots, A_{iM}$ , respectively, and the values of the other elements are zero to a first gain circuit 30 as a first sound source signal (sound source vector).

[0021] A second pulse position vector  $Q = (Q_1, Q_2, \dots, Q_M)$  output by the second pulse position generating circuit 611 and a second pulse amplitude vector  $B_j = (B_{i1}, B_{i2}, \dots, B_{iM})$  output by the second pulse amplitude generating circuit 121 enter the second sound source generating circuit 21. The second sound source generating circuit 21 outputs an N-dimensional vector for which the values of the  $Q_1$ st,  $Q_2$ nd, ...,  $Q_M$ th elements are  $B_{i1}, B_{i2}, \dots, B_{iM}$ , respectively, and the values of the other elements are zero to a second gain circuit 31 as a second sound source signal.

[0022] The first gain circuit 30 has a table in which

gain values have been stored. The index output by the minimizing circuit 670 and the first sound source vector output by the first sound source generating circuit 20 enter the first gain circuit 30, which proceeds to read a first gain corresponding to the index out of the table, multiply the first gain by the first sound source vector to thereby generate a third sound source vector, and output the generated third sound source vector to a first higher-order linear prediction filter 130.

[0023] The second gain circuit 31 has a table in which gain values have been stored. The index output by the minimizing circuit 670 and the second sound source vector output by the second sound source generating circuit 21 enter the second gain circuit 31, which proceeds to read a second gain corresponding to the index out of the table, multiply the second gain by the second sound source vector to thereby generate a fourth sound source vector, and output the generated fourth sound source vector to a second higher-order linear prediction filter 131.

[0024] A third higher-order linear prediction coefficient output by a higher-order linear prediction coefficient calculation circuit 180 and a third sound source vector output by the first gain circuit 30 enter the first higher-order linear prediction filter 130. The filter thus set to the third higher-order linear prediction coefficient is driven by the third sound source vector, whereby a first excitation vector is obtained. The first excitation vector is output to a first band-pass filter 135.

[0025] A fourth higher-order linear prediction coefficient output by the higher-order linear prediction coefficient calculation circuit 180 and a fourth sound source vector output by the second gain circuit 31 enter the second higher-order linear prediction filter 131. The filter thus set to the fourth higher-order linear prediction coefficient is driven by the fourth sound source vector, whereby a second excitation vector is obtained. The second excitation vector is output to a second band-pass filter 136.

[0026] The first excitation vector output by the first higher-order linear prediction filter 130 enters the first band-pass filter 135. The first excitation vector has its band limited by the filter 135, whereby a third excitation vector is obtained. The first band-pass filter 135 outputs the third excitation vector to an adder 40.

[0027] The second excitation vector output by the second higher-order linear prediction filter 131 enters the second band-pass filter 136. The second excitation vector has its band limited by the filter 136, whereby a fourth excitation vector is obtained. The fourth excitation vector is output to the adder 40.

[0028] The adder 40 adds the inputs applied thereto, namely the third excitation vector output by the first band-pass filter 135 and the fourth excitation vector output by the second band-pass filter 136, and outputs a fifth excitation vector, which is the sum of the third and fourth excitation vectors, to the linear prediction filter 150.

[0029] The linear prediction filter 150 has a table in which quantized values of linear prediction coefficients have been stored. The fifth excitation vector output by the adder 40 and an index corresponding to a quantized value of a linear prediction coefficient output by the first linear prediction coefficient calculation circuit 140 enter the linear prediction filter 150. The quantized value of the linear prediction coefficient corresponding to this index is read out of this table and the filter thus set to this quantized linear prediction coefficient is driven by the fifth excitation vector, whereby a reconstructed signal (reconstructed vector) is obtained. This vector is output to a subtractor 50 and to the higher-order linear prediction coefficient calculation circuit 180.

[0030] The reconstructed vector output by the linear prediction filter 150 enters the higher-order linear prediction coefficient calculation circuit 180, which proceeds to calculate the third higher-order linear prediction coefficient and the fourth higher-order linear prediction coefficient. The third higher-order linear prediction coefficient is output to the first higher-order linear prediction filter 130, and the fourth higher-order linear prediction coefficient is output to the second higher-order linear prediction filter 131. The details of construction of the higher-order linear prediction coefficient calculation circuit 180 will be described later.

[0031] The input vector enters the subtractor 50 via the input terminal 10, and the reconstructed vector output by the linear prediction filter 150 also enters the subtractor 50. The subtractor 50 calculates the difference between these two inputs. The subtractor 50 outputs a difference vector, which is the difference between the input vector and the reconstructed vector, to the weighting filter 160.

[0032] The difference vector output by the subtractor 50 and the linear prediction coefficient output by the first linear prediction coefficient calculation circuit 140 enter the weighting filter 160. The latter uses this linear prediction coefficient to produce a weighting filter corresponding to the characteristic of the human sense of hearing and drives this weighting filter by the difference vector, whereby there is obtained a weighted difference vector. The weighted difference vector is output to the minimizing circuit 670. For a discussion of a weighting filter, see Reference 1.

[0033] Weighted difference vectors output by the weighting filter 160 successively enter the minimizing circuit 670, which proceeds to calculate the norms.

[0034] Indices corresponding to all values of the elements of the first pulse position vector in the first pulse position generating circuit 610 are output successively from the minimizing circuit 670 to the first pulse position generating circuit 610. Indices corresponding to all values of the elements of the second pulse position vector in the second pulse position generating circuit 611 are output successively from the minimizing circuit 670 to the second pulse position generating circuit 611. Indices corresponding to all first pulse amplitude vectors

5 that have been stored in the first pulse amplitude generating circuit 120 are output successively from the minimizing circuit 670 to the first pulse amplitude generating circuit 120. Indices corresponding to all second pulse amplitude vectors that have been stored in the second pulse amplitude generating circuit 121 are output successively from the minimizing circuit 670 to the second pulse amplitude generating circuit 121. Indices corresponding to all first gains that have been stored in the first gain circuit 30 are output successively from the minimizing circuit 670 to the first gain circuit 30. Indices corresponding to all second gains that have been stored in the second gain circuit 31 are output successively from the minimizing circuit 670 to the second gain circuit 31.

10 Further, the minimizing circuit 670 selects the value of each element in the first pulse position vector, the value of each element in the second pulse position vector, the first pulse amplitude vector, the second pulse amplitude vector and the first gain and second gain that will result in the minimum norm and outputs the indices corresponding to these to the code output circuit 690.

[0035] With regard to a method of obtaining the position of each pulse that is an element of a pulse position vector as well as the amplitude value of each pulse that is an element of a pulse amplitude vector, see Reference 4, by way of example.

[0036] The index corresponding to the quantized value of the linear prediction coefficient output by the first linear prediction coefficient calculation circuit 140 enters the code output circuit 690 and so do the indices corresponding to the value of each element in the first pulse position vector, the value of each element in the second pulse position vector, the first pulse amplitude vector, the second pulse amplitude vector and the first gain and second gain. The code output circuit 690 converts these indices to a bit-sequence code and outputs the code via an output terminal 60.

[0037] The higher-order linear prediction coefficient calculation circuit 180 will now be described with reference to Fig. 11.

[0038] As shown in Fig. 11, the reconstructed vector output by the linear prediction filter 150 enters a second linear prediction coefficient calculation circuit 910 via an input terminal 900. The second linear prediction coefficient calculation circuit 910 subjects this reconstructed vector to linear prediction analysis, obtains a linear prediction coefficient and outputs this coefficient to a residual signal calculation circuit 920 as a second linear prediction coefficient.

[0039] The second linear prediction coefficient output by the second linear prediction coefficient calculation circuit 910 and the reconstructed vector output by the linear prediction filter 150 enter the residual signal calculation circuit 920, which proceeds to use a filter, in which the second linear prediction coefficient has been set, to subject the reconstructed vector to inverse filtering, whereby a first residual vector is obtained. The first residual vector is output to an FFT (Fast-Fourier Trans-

form) circuit 930.

[0040] The FFT circuit 930, to which the first residual vector output by the residual signal calculation circuit 920 is applied, subjects this vector to a Fourier transform and outputs the Fourier coefficients thus obtained to a band splitting circuit 940.

[0041] The band splitting circuit 940, to which the Fourier coefficients output by the FFT circuit 930 are applied, equally partitions these Fourier coefficients into high-and low-frequency regions, thereby obtaining low-frequency Fourier coefficients and high-frequency Fourier coefficients. The low-frequency coefficients are output to a first downsampling circuit 950 and the high-frequency coefficients are output to a second downsampling circuit 951.

[0042] The first downsampling circuit 950 downsamples the low-frequency Fourier coefficients output by the band splitting circuit 940. Specifically, the first downsampling circuit 950 removes bands corresponding to high frequency in the low-frequency Fourier coefficients and generates first Fourier coefficients the band whereof is half the full band. The first Fourier coefficients are output to a first inverse FFT circuit 960.

[0043] The second downsampling circuit 951 downsamples the high-frequency Fourier coefficients output by the band splitting circuit 940. Specifically, the second downsampling circuit 951 removes bands corresponding to low frequency in the high-frequency Fourier coefficients and loops back the high-frequency coefficients to the low-frequency side, thereby generating second Fourier coefficients the band whereof is half the full band. The second Fourier coefficients are output to a second inverse FFT circuit 961.

[0044] The first Fourier coefficients output by the first downsampling circuit 950 enter the first inverse FFT circuit 960, which proceeds to subject these coefficients to an inverse FFT, thereby obtaining a second residual vector that is output to a first higher-order linear prediction coefficient calculation circuit 970.

[0045] The second Fourier coefficients output by the second downsampling circuit 951 enter the second inverse FFT circuit 961, which proceeds to subject these coefficients to an inverse FFT, thereby obtaining a third residual vector that is output to a second higher-order linear prediction coefficient calculation circuit 971.

[0046] The second residual vector output by the first inverse FFT circuit 960 enters the first higher-order linear prediction coefficient calculation circuit 970, which proceeds to subject the second residual vector to higher-order linear prediction analysis, thereby obtaining the first higher-order linear prediction coefficient. This is output to a first upsampling circuit 980.

[0047] The third residual vector output by the second inverse FFT circuit 961 enters the second higher-order linear prediction coefficient calculation circuit 971, which proceeds to subject the third residual vector to higher-order linear prediction analysis, thereby obtaining the second higher-order linear prediction coefficient.

This is output to a second upsampling circuit 981.

[0048] The first higher-order linear prediction coefficient output by the first higher-order linear prediction coefficient calculation circuit 970 enters the first upsampling circuit 980. By inserting zeros in alternation with the first higher-order linear prediction coefficient, the first upsampling circuit 980 obtains an upsampled prediction coefficient. This is output as the third higher-order linear prediction coefficient to the first higher-order linear prediction filter 130 via an output terminal 901.

[0049] The second higher-order linear prediction coefficient output by the second higher-order linear prediction coefficient calculation circuit 971 enters the second upsampling circuit 981. By inserting zeros in alternation with the second higher-order linear prediction coefficient, the second upsampling circuit 981 obtains an upsampled prediction coefficient. This is output as the fourth higher-order linear prediction coefficient to the second higher-order linear prediction filter 131 via an output terminal 902.

[0050] Fig. 12 is a block diagram showing an example of the construction of an apparatus for decoding speech and music according to the prior art. Components in Fig. 12 identical with or equivalent to those of Fig. 10 are designated by like reference characters.

[0051] As shown in Fig. 12, a code in the form of a bit sequence enters from an input terminal 200. A code input circuit 720 converts the bit-sequence code that has entered from the input terminal 200 to an index.

[0052] The code input circuit 720 outputs an index corresponding to each element in the first pulse position vector to a first pulse position generating circuit 710, outputs an index corresponding to each element in the second pulse position vector to a second pulse position generating circuit 711, outputs an index corresponding to the first pulse amplitude vector to the first pulse amplitude generating circuit 120, outputs an index corresponding to the second pulse amplitude vector to the second pulse amplitude generating circuit 121, outputs an index corresponding to the first gain to the first gain circuit 30, outputs an index corresponding to the second gain to the second gain circuit 31, and outputs an index corresponding to the quantized value of a linear prediction coefficient to the linear prediction filter 150.

[0053] The index output by the code input circuit 720 enters the first pulse position generating circuit 710, which proceeds to generate the first pulse position vector using the position of each pulse specified by the index and output the vector to the first sound source generating circuit 20.

[0054] The first pulse amplitude generating circuit 120 has a table in which M-dimensional vectors  $A'_j$ ,  $j = 1, \dots, N_A$  have been stored. The index output by the code input circuit 720 enters the first pulse amplitude generating circuit 120, which proceeds to read an M-dimensional vector  $A'_j$  corresponding to this index out of the above-mentioned table and to output this vector to the

first sound source generating circuit 20 as a first pulse amplitude vector.

[0055] The index output by the code input circuit 720 enters the second pulse position generating circuit 711, which proceeds to generate the second pulse position vector using the position of each pulse specified by the index and output the vector to the second sound source generating circuit 21.

[0056] The second pulse amplitude generating circuit 121 has a table in which M-dimensional vectors  $B_j$ ,  $j = 1, \dots, N_B$  have been stored. The index output by the code input circuit 720 enters the second pulse amplitude generating circuit 121, which proceeds to read an M-dimensional vector  $B_j$  corresponding to this index out of the above-mentioned table and to output this vector to the second sound source generating circuit 21 as a second pulse amplitude vector.

[0057] The first pulse position vector  $P^- = (P_1, P_2, \dots, P_M)$  output by the first pulse position generating circuit 710 and the first pulse amplitude vector  $A_i^- = (A_{i1}, A_{i2}, \dots, A_{iM})$  output by the first pulse amplitude generating circuit 120 enter the first sound source generating circuit 20. The first sound source generating circuit 20 outputs an N-dimensional vector for which the values of the  $P_1$ st,  $P_2$ nd, ...,  $P_M$ th elements are  $A_{i1}, A_{i2}, \dots, A_{iM}$ , respectively, and the values of the other elements are zero to the first gain circuit 30 as a first sound source signal vector.

[0058] The second pulse position vector  $Q^- = (Q_1, Q_2, \dots, Q_M)$  output by the second pulse position generating circuit 711 and the second pulse amplitude vector  $B_i^- = (B_{i1}, B_{i2}, \dots, B_{iM})$  output by the second pulse amplitude generating circuit 121 enter the second sound source generating circuit 21. The second sound source generating circuit 21 outputs an N-dimensional vector for which the values of the  $Q_1$ st,  $Q_2$ nd, ...,  $Q_M$ th elements are  $B_{i1}, B_{i2}, \dots, B_{iM}$ , respectively, and the values of the other elements are zero to the second gain circuit 31 as a second sound source signal.

[0059] The first gain circuit 30 has a table in which gain values have been stored. The index output by the code input circuit 720 and the first sound source vector output by the first sound source generating circuit 20 enter the first gain circuit 30, which proceeds to read a first gain corresponding to the index out of the table, multiply the first gain by the first sound source vector to thereby generate a third sound source vector and output the generated third sound source vector to the first higher-order linear prediction filter 130.

[0060] The first gain circuit 31 has a table in which gain values have been stored. The index output by the code input circuit 720 and the second sound source vector output by the second sound source generating circuit 21 enter the second gain circuit 31, which proceeds to read a second gain corresponding to the index out of the table, multiply the second gain by the second sound source vector to thereby generate a fourth sound source vector and output the generated fourth sound

source vector to a second higher-order linear prediction filter 131.

[0061] The third higher-order linear prediction coefficient output by the higher-order linear prediction coefficient calculation circuit 180 and the third sound source vector output by the first gain circuit 30 enter the first higher-order linear prediction filter 130. The filter thus set to the third higher-order linear prediction coefficient is driven by the third sound source vector, whereby a first excitation vector is obtained. The first excitation vector is output to the first band-pass filter 135.

[0062] The fourth higher-order linear prediction coefficient output by the higher-order linear prediction coefficient calculation circuit 180 and the fourth sound source vector output by the second gain circuit 31 enter the second higher-order linear prediction filter 131. The filter thus set to the fourth higher-order linear prediction coefficient is driven by the fourth sound source vector, whereby a second excitation vector is obtained. The second excitation vector is output to the second band-pass filter 136.

[0063] The first excitation vector output by the first higher-order linear prediction filter 130 enters the first band-pass filter 135. The first excitation vector has its band limited by the filter 135, whereby a third excitation vector is obtained. The first band-pass filter 135 outputs the third excitation vector to the adder 40.

[0064] The second excitation vector output by the second higher-order linear prediction filter 131 enters the second band-pass filter 136. The second excitation vector has its band limited by the filter 136, whereby a fourth excitation vector is obtained. The fourth excitation vector is output to the adder 40.

[0065] The adder 40 adds the inputs applied thereto, namely the third excitation vector output by the first band-pass filter 135 and the fourth excitation vector output by the second band-pass filter 136, and outputs a fifth excitation vector, which is the sum of the third and fourth excitation vectors, to the linear prediction filter 150.

[0066] The linear prediction filter 150 has a table in which quantized values of linear prediction coefficients have been stored. The fifth excitation vector output by the adder 40 and an index corresponding to a quantized value of a linear prediction coefficient output by the code input circuit 720 enter the linear prediction filter 150. The latter reads the quantized value of the linear prediction coefficient corresponding to this index out of the table and drives the filter thus set to this quantized linear prediction coefficient by the fifth excitation vector, whereby a reconstructed vector is obtained.

[0067] The reconstructed vector obtained is output to an output terminal 201 and to the higher-order linear prediction coefficient calculation circuit 180.

[0068] The reconstructed vector output by the linear prediction filter 150 enters the higher-order linear prediction coefficient calculation circuit 180, which proceeds to calculate the third higher-order linear

prediction coefficient and the fourth higher-order linear prediction coefficient. The third higher-order linear prediction is output to the first higher-order linear prediction filter 130, and the fourth higher-order linear prediction coefficient is output to the second higher-order linear prediction filter 131.

[0069] The reconstructed vector calculated by the linear prediction filter 150 is output via the output terminal 201.

[0070] In the course of investigations toward the present invention, the following problem has been encountered. Namely, a problem with the conventional apparatus for encoding and decoding speech and musical signals by the above-described band splitting technique is that a large number of bits is required to encode the sound source signals.

[0071] The reason for this is that the sound source signals are encoded independently in each band without taking into consideration the correlation between bands of the input signals.

[0072] Accordingly, an object of the present invention is to provide an apparatus for encoding and decoding speech and musical signals, wherein the sound source signal of each band can be encoded using a small number of bits.

[0073] Another object of the present invention is to provide an apparatus for encoding or decoding speech and musical (i.e., sound) signals with simplified structure and/or high efficiency. Further objects of the present invention will become apparent in the entire disclosure. Generally, the present invention contemplates to utilize the correlation between bands of the input signals upon encoding/decoding in such a fashion to reduce the entire bit number.

[0074] According to a first aspect of the present invention, the foregoing object is attained by providing a speech and musical signal encoding apparatus which, when encoding an input signal upon splitting the input signal into a plurality of bands, generates a reconstructed signal using a multi pulse sound source signal that corresponds to each band, wherein a position obtained by shifting the position of each pulse which defines the multipulse signal in the band(s) is used when defining a multipulse signal in the other band(s).

[0075] According to a second aspect of the present invention, the foregoing object is attained by providing a speech and musical signal decoding apparatus for generating a reconstructed signal using a multi pulse sound source signal corresponding to each of a plurality of bands, wherein a position obtained by shifting the position of each pulse which defines the multipulse signal in the band(s) is used when defining a multipulse signal in the other band(s).

[0076] According to a third aspect of the present invention, the foregoing object is attained by providing a speech and musical signal encoding apparatus which, when encoding an input signal upon splitting the input signal into a plurality of bands, generates a recon-

structed signal by exciting a synthesis filter by a full-band sound source signal, which is obtained by summing, over all bands, multipulse sound source signals corresponding to respective ones of the plurality of bands, wherein a position obtained by shifting the position of each pulse which defines the multipulse signal in the band(s) is used when defining a multipulse signal in the other band(s).

[0077] According to a fourth aspect of the present invention, the foregoing object is attained by providing a speech and musical signal decoding apparatus for generating a reconstructed signal by exciting a synthesis filter by a full-band sound source signal, which is obtained by summing, over all bands, multi pulse sound source signals corresponding to respective ones of a plurality of bands, wherein a position obtained by shifting the position of each pulse which defines the multi pulse signal in the band(s) is used when defining a multipulse signal in the other band(s).

[0078] According to a fifth aspect of the present invention, the foregoing object is attained by providing a speech and musical signal encoding apparatus which, when encoding an input signal upon splitting the input signal into a plurality of bands, generates a reconstructed signal by exciting a synthesis filter by a full-band sound source signal, which is obtained by summing, over all bands, signals obtained by exciting a higher-order linear prediction filter, which represents a microspectrum relating to the input signal of each band, by a multipulse sound source signal corresponding to each band, wherein a position obtained by shifting the position of each pulse which defines the multipulse signal in the band(s) is used when defining a multipulse signal in the other band(s).

[0079] According to a sixth aspect of the present invention, the foregoing object is attained by providing a speech and musical signal decoding apparatus for generating a reconstructed signal by exciting a synthesis filter by a full-band sound source signal, which is obtained by summing, over all bands, signals obtained by exciting a higher-order linear prediction filter, which represents a microspectrum relating to an input signal of each of a plurality of bands, by a multipulse sound source signal corresponding to each band, wherein a position obtained by shifting the position of each pulse which defines the multipulse signal in the band(s) is used when defining a multipulse signal in the other band(s).

[0080] According to a seventh aspect of the present invention, the foregoing object is attained by providing a speech and musical signal encoding apparatus which, when encoding an input signal upon splitting the input signal into a plurality of bands, generates a reconstructed signal by exciting a synthesis filter by a full-band sound source signal, which is obtained by summing, over all bands, signals obtained by exciting a higher-order linear prediction filter, which represents a microspectrum relating to the input signal of each band, by a multipulse sound source signal corresponding to

each band, wherein a residual signal is found by inverse filtering of the reconstructed signal using a linear prediction filter for which linear prediction coefficients obtained from the reconstructed signal have been decided, conversion coefficients obtained by converting the residual signal are split into bands, and the higher-order linear prediction filter uses coefficients obtained from a residual signal of each band generated in each band by back-converting the conversion coefficients that have been split into the bands.

[0081] According to an eighth aspect of the present invention, the foregoing object is attained by providing a speech and musical signal decoding apparatus for generating a reconstructed signal by exciting a synthesis filter by a full-band sound source signal, which is obtained by summing, over all bands, signals obtained by exciting a higher-order linear prediction filter, which represents a microspectrum relating to an input signal of each of a plurality of bands, by a multipulse sound source signal corresponding to each band, wherein a residual signal is found by inverse filtering of the reconstructed signal using a linear prediction filter for which linear prediction coefficients obtained from the reconstructed signal have been decided, conversion coefficients obtained by converting the residual signal are split into bands, and the higher-order linear prediction filter uses coefficients obtained from a residual signal of each band generated in each band by back-converting the conversion coefficients that have been split into the bands.

[0082] According to a ninth aspect of the present invention, in the fifth aspect of the invention a residual signal is found by inverse filtering of the reconstructed signal using a linear prediction filter for which linear prediction coefficients obtained from the reconstructed signal have been decided, conversion coefficients obtained by converting the residual signal are split into bands, and the higher-order linear prediction filter uses coefficients obtained from a residual signal of each band generated in each band by back-converting the conversion coefficients that have been split into the bands.

[0083] According to a tenth aspect of the present invention, in the sixth aspect of the invention a residual signal is found by inverse filtering of the reconstructed signal using a linear prediction filter for which linear prediction coefficients obtained from the reconstructed signal have been decided, conversion coefficients obtained by converting the residual signal are split into bands, and the higher-order linear prediction filter uses coefficients obtained from a residual signal of each band generated in each band by back-converting the conversion coefficients that have been split into the bands.

[0084] Other features and advantages of the present invention will be apparent from the following description taken in conjunction with the accompanying drawings, in which like reference characters designate the same or similar parts throughout the figures thereof.

5

10

15

20

25

30

35

40

45

50

55

Fig. 1 is a block diagram illustrating the construction of a first embodiment of an apparatus for encoding speech and musical signals according to the present invention;

Fig. 2 is a block diagram illustrating the construction of a first embodiment of an apparatus for decoding speech and musical signals according to the present invention;

Fig. 3 is a block diagram illustrating the construction of a second embodiment of an apparatus for encoding speech and musical signals according to the present invention;

Fig. 4 is a block diagram illustrating the construction of a second embodiment of an apparatus for decoding speech and musical signals according to the present invention;

Fig. 5 is a block diagram illustrating the construction of a third embodiment of an apparatus for encoding speech and musical signals according to the present invention;

Fig. 6 is a block diagram illustrating the construction of a higher-order linear prediction coefficient calculation circuit according to the third embodiment;

Fig. 7 is a block diagram illustrating the construction of a third embodiment of an apparatus for decoding speech and musical signals according to the present invention;

Fig. 8 is a block diagram illustrating the construction of a fourth embodiment of an apparatus for encoding speech and musical signals according to the present invention;

Fig. 9 is a block diagram illustrating the construction of a fourth embodiment of an apparatus for decoding speech and musical signals according to the present invention;

Fig. 10 is a block diagram illustrating the construction of an apparatus for encoding speech and musical signals according to the prior art prior art;

Fig. 11 is a block diagram illustrating the construction of a higher-order linear prediction coefficient calculation circuit according to the prior art; and

Fig. 12 is a block diagram illustrating the construction of a fourth embodiment of an apparatus for decoding speech and musical signals according to the prior art.

[0085] Preferred modes of practicing the present invention will now be described. An apparatus for encoding speech and musical signals according to the present invention in a first preferred mode thereof generates a reconstructed signal using a multipulse sound source signal that corresponds to each of a plurality of bands when a speech input signal is encoded upon being split into a plurality of bands, wherein some of the information possessed by a sound source signal encoded in a certain band is used to encode a sound source signal in another band. More specifically, the encoding apparatus has means (a first pulse position

generating circuit 110, a second pulse position generating circuit 111 and a minimizing circuit 170 shown in Fig. 1) for using a position obtained by shifting the position of each pulse, which defines the multipulse signal in the band or bands, when a multipulse signal in the other band(s) is defined.

[0086] More specifically, in regard to a case where the number of bands is two, for example, an index output by the minimizing circuit 170 in Fig. 1 and a first pulse position vector  $P^- = (P_1, P_2, \dots, P_M)$  output by the minimizing circuit 170 enter the second pulse position generating circuit 111. The latter revises the first pulse position vector using a pulse position revision quantity  $d^-_i = (d_{i1}, d_{i2}, \dots, d_{iM})$  specified by the index and outputs the revised vector to the second sound source generating circuit 21 in Fig. 1 as a second pulse position vector  $P^- = (P_1+d_{i1}, P_2+d_{i2}, \dots, P_M+d_{iM})$ .

[0087] An apparatus for decoding speech and musical signals according to the present invention in the first preferred mode thereof uses some of the information possessed by a sound source signal decoded in certain band or bands to decode a sound source signal in another band or the other bands.

[0088] More specifically, the decoding apparatus has means (a first pulse position generating circuit 210, a second pulse position generating circuit 211 and a code input circuit 220 shown in Fig. 2) for using a position obtained by shifting the position of each pulse, which defines the multi pulse signal in the band, when a multi-pulse signal in another band is defined.

[0089] An apparatus for encoding speech and musical signals according to the present invention in a second preferred mode thereof generates a reconstructed signal by exciting a synthesis filter by a full-band sound source signal, which is obtained by summing, over all bands, multipulse sound source signals corresponding to respective ones of the plurality of bands. More specifically, the encoding apparatus has means (110, 111, 170 in Fig. 1) for using a position obtained by shifting the position of each pulse, which defines the multipulse signal in the band(s), when a multipulse signal in the other band(s) is defined, means (adder 40 in Fig. 1) for obtaining the full-band sound source signal by summing, over all bands, multipulse sound source signals corresponding to respective ones of the bands, and means (linear prediction filter 150 in Fig. 1) for generating the reconstructed signal by exciting the synthesis filter by the full-band sound source signal.

[0090] An apparatus for decoding speech and musical signals according to the present invention in the second preferred mode thereof generates a reconstructed signal by exciting a synthesis filter by a full-band sound source signal, which is obtained by summing, over all bands, multipulse sound source signals corresponding to respective ones of the plurality of bands. More specifically, the decoding apparatus has means (210, 211 and 220 in Fig. 2) for using a position obtained by shifting the position of each pulse, which defines the multipulse sig-

nal in the band(s), when a multipulse signal in the other band(s) is defined; means (adder 40 in Fig. 2) for obtaining the full-band sound source signal by summing, over all bands, multipulse sound source signals corresponding to respective ones of the bands; and means (linear prediction filter 150 in Fig. 1) for generating the reconstructed signal by exciting the synthesis filter by the full-band sound source signal.

[0091] An apparatus for encoding speech and musical signals according to the present invention in a third preferred mode thereof generates a reconstructed signal by exciting a synthesis filter by a full-band sound source signal, which is obtained by summing, over all bands, signals obtained by exciting a higher-order linear prediction filter, which represents a microspectrum relating to the input signal of each band, by a multi pulse sound source signal corresponding to each band. More specifically, the encoding apparatus has means (the first pulse position generating circuit 110, second pulse position generating circuit 111 and minimizing circuit 170 shown in Fig. 1) for using a position obtained by shifting the position of each pulse, which defines the multipulse signal in the band(s), when a multipulse signal in the other band(s) is defined; means (first and second higher-order linear prediction filters 130, 131 in Fig. 3) for exciting the higher-order linear prediction filter by the multi-pulse sound source signal corresponding to each band; means (adder 40 in Fig. 3) for obtaining the full-band sound source signal by summing, over all bands, signals obtained by exciting the higher-order linear prediction filter; and means (linear prediction filter 150 in Fig. 3) for generating the reconstructed signal by exciting the synthesis filter by the full-band sound source signal.

[0092] An apparatus for decoding speech and musical signals according to the present invention in the third preferred mode thereof generates a reconstructed signal by exciting a synthesis filter by a full-band sound source signal, which is obtained by summing, over all bands, signals obtained by exciting a higher-order linear prediction filter, which represents a microspectrum relating to the input signal of each band, by a multipulse sound source signal corresponding to each band. More specifically, the decoding apparatus has means (first pulse position generating circuit 210, second pulse position generating circuit 211 and code input circuit 220 shown in Fig. 4) for using a position obtained by shifting the position of each pulse, which defines the multipulse signal in the band(s), when a multipulse signal in the other band(s) is defined; means (first and second higher-order linear prediction filters 130, 131 in Fig. 4) for exciting the higher-order linear prediction filter by the multipulse sound source signal corresponding to each band; means (adder 40 in Fig. 4) for obtaining the full-band sound source signal by summing, over all bands, signals obtained by exciting the higher-order linear prediction filter; and means (linear prediction filter 150 in Fig. 4) for generating the reconstructed signal by exciting the synthesis filter by the full-band sound

source signal.

[0093] In a fourth preferred mode of the present invention, the apparatus for encoding speech and musical signals of the third mode is characterized in that a higher-order linear prediction calculation circuit is implemented by a simple arrangement. More specifically, the encoding apparatus has means (second linear prediction coefficient calculation circuit 910 and residual signal calculation circuit 920 in Fig. 6) for obtaining a residual signal by inverse filtering of the reconstructed signal using a linear prediction filter for which linear prediction coefficients obtained from the reconstructed signal have been decided and set; means (FFT circuit 930 and band splitting circuit 540 in Fig. 6) for splitting, into bands, conversion coefficients obtained by converting the residual signal; and means (first zero fill circuit 550, second zero fill circuit 551, first inverse FFT circuit 560, second inverse FFT circuit 561, first higher-order linear prediction coefficient calculation circuit 570 and second higher-order linear prediction coefficient calculation circuit 571 in Fig. 6) for outputting, to the higher-order linear prediction filter, coefficients obtained from a residual signal of each band generated in each band by back-converting the conversion coefficients that have been split into the bands.

[0094] In a fourth preferred mode of the present invention, the apparatus for decoding speech and musical signals of the third mode is characterized in that a higher-order linear prediction calculation circuit is implemented by a simple arrangement. More specifically, the encoding apparatus has means (910, 920 in Fig. 6) for obtaining a residual signal by inverse filtering of the reconstructed signal using a linear prediction filter for which linear prediction coefficients obtained from the reconstructed signal have been decided; means (930, 540 in Fig. 6) for splitting, into bands, conversion coefficients obtained by converting the residual signal; and means (550, 551, 560, 561, 570, 571 in Fig. 6) for outputting, to the higher-order linear prediction filter, coefficients obtained from a residual signal of each band generated in each band by back-converting the conversion coefficients that have been split into the bands.

[0095] In a fifth preferred mode of the present invention, the apparatus for encoding speech and musical signals of the fourth mode is further characterized in that the sound source signal of each band is encoded independently. More specifically, the encoding apparatus has means (first pulse position generating circuit 510, second pulse position generating circuit 511 and minimizing circuit 670 in Fig. 8) for separately obtaining, in each band, the position of each pulse defining the multipulse signal.

[0096] In the fifth preferred mode of the present invention, the apparatus for decoding speech and musical signals of the fourth mode is further characterized in that the sound source signal of each band is decoded independently. More specifically, the decoding apparatus has means (first pulse position generating circuit

710, second pulse position generating circuit 711 and code input circuit 720 in Fig. 9) for separately(individually) obtaining, in each band, the position of each pulse defining the multipulse signal.

5 [0097] In the modes of the present invention described above, some of the information possessed by a sound source signal that has been encoded in a certain band or bands is used to encode a sound source signal in the other band or bands. That is, encoding is performed taking into account the correlation between bands possessed by the input signal. More specifically, the position of each pulse obtained by uniformly shifting the positions of the pulses obtained when a multipulse sound source signal is encoded in a first band is used when encoding a sound source signal in a second band.

10 [0098] As a consequence, in relation to the sound source signal in the second band, the number of bits necessary in the conventional method to separately represent the position of each pulse is reduced to a number of bits necessary solely for representing the amount of shift.

15 [0099] As a result, it is possible to reduce the number of bits needed to encode the sound source signal in the second band.

20 [0100] Embodiments of the present invention will now be described with reference to the drawings in order to explain further the modes of the invention set forth above.

25 30 [First Embodiment]

[0101] Fig. 1 is a block diagram illustrating the construction of a first embodiment of an apparatus for encoding speech and musical signals according to the present invention. Here it is assumed for the sake of simplicity that the number of bands is two.

[0102] As shown in Fig. 1, an input vector enters from the input terminal 10. The first linear prediction coefficient calculation circuit 140 receives the input vector as an input from the input terminal 10 and this circuit subjects the input vector to linear prediction analysis, obtains a linear prediction coefficient and quantizes the coefficient. The first linear prediction coefficient calculation circuit 140 outputs the linear prediction coefficient to the weighting filter 160 and outputs an index, which corresponds to a quantized value of the linear prediction coefficient, to the linear prediction filter 150 and to a code output circuit 190.

[0103] The first pulse position generating circuit 110 receives as an input an index that is output by the minimizing circuit 170, generates a first pulse position vector  $P^*$  using the position of each pulse specified by the index and outputs this vector to the first sound source generating circuit 20 and to the second pulse position generating circuit 111.

[0104] Let  $M$  represent the number of pulses and let  $P_1, P_2, \dots, P_M$  represent the positions of the pulses. The vector  $P^*$ , therefore, is written as follows:

$$P^* = (P_1, P_2, \dots, P_M)$$

[0105] The first pulse amplitude generating circuit 120 has a table in which M-dimensional vectors  $A_i^*$ ,  $j = 1, \dots, N_A$  have been stored, where  $N_A$  represents the size of the table. The index output by the minimizing circuit 170 enters the first pulse amplitude generating circuit 120, which proceeds to read an M-dimensional vector  $A_i^*$  corresponding to this index out of the above-mentioned table and to output this vector to the first sound source generating circuit 20 as a first pulse amplitude vector.

[0106] Letting  $A_{i1}, A_{i2}, \dots, A_{iM}$  represent the amplitude values of the pulses, we have  $A_i^* = (A_{i1}, A_{i2}, \dots, A_{iM})$ .

[0107] The second pulse position generating circuit 111 receives as inputs the index that is output by the minimizing circuit 170 and the first pulse position vector  $P^* = (P_1, P_2, \dots, P_M)$  output by the first pulse position generating circuit 110, revises the first pulse position vector using the pulse position revision quantity  $d_i^* = (d_{i1}, d_{i2}, \dots, d_{iM})$  specified by the index and outputs the revised vector to the second sound source generating circuit 21 as a second pulse position vector  $Q^* = (P_1+d_{i1}, P_2+d_{i2}, \dots, P_M+d_{iM})$ .

[0108] The second pulse amplitude generating circuit 121 has a table in which M-dimensional vectors  $B_j^*$ ,  $j = 1, \dots, N_B$  have been stored, where  $N_B$  represents the size of the table.

[0109] The index output by the minimizing circuit 170 enters the second pulse amplitude generating circuit 121, which proceeds to read an M-dimensional vector  $B_i^*$  corresponding to this index out of the above-mentioned table and to output this vector to the second sound source generating circuit 21 as a second pulse amplitude vector.

[0110] The first pulse position vector  $P^* = (P_1, P_2, \dots, P_M)$  output by the first pulse position generating circuit 110 and the first pulse amplitude vector  $A_i^* = (A_{i1}, A_{i2}, \dots, A_{iM})$  output by the first pulse amplitude generating circuit 120 enter the first sound source generating circuit 20. The first sound source generating circuit 20 outputs an N-dimensional vector for which the values of the  $P_1$ st,  $P_2$ nd, ...,  $P_M$ th elements are  $A_{i1}, A_{i2}, \dots, A_{iM}$ , respectively, and the values of the other elements are zero to the first gain circuit 30 as a first sound source vector.

[0111] A second pulse position vector  $Q^* = (Q_1^*, Q_2^*, \dots, Q_M^*)$  output by the second pulse position generating circuit 111 and a second pulse amplitude vector  $B_i^* = (B_{i1}, B_{i2}, \dots, B_{iM})$  output by the second pulse amplitude generating circuit 121 enter the second sound source generating circuit 21. The second sound source generating circuit 21 outputs an N-dimensional vector for which the values of the  $Q_1^*$ st,  $Q_2^*$ nd, ...,  $Q_M^*$ th elements are  $B_{i1}, B_{i2}, \dots, B_{iM}$ , respectively, and the values of the other elements are zero to a second gain circuit 31 as a second sound source vector.

[0112] The first gain circuit 30 has a table in which gain values have been stored. The index output by the

minimizing circuit 170 and the first sound source vector output by the first sound source generating circuit 20 enter the first gain circuit 30, which proceeds to read a first gain corresponding to the index out of the table, multiply the first gain by the first sound source vector to thereby generate a third sound source vector, and output the generated third sound source vector to the first band-pass filter 135.

[0113] The second gain circuit 31 has a table in which gain values have been stored. The index output by the minimizing circuit 170 and the second sound source vector output by the second sound source generating circuit 21 enter the second gain circuit 31, which proceeds to read a second gain corresponding to the index out of the table, multiply the second gain by the second sound source vector to thereby generate a fourth sound source vector, and output the generated fourth sound source vector to the second band-pass filter 136.

[0114] The third sound source vector output by the first gain circuit 30 enters the first band-pass filter 135. The third sound source vector has its band limited by the filter 135, whereby a fifth sound source vector is obtained. The first band-pass filter 135 outputs the fifth sound source vector to the adder 40.

[0115] The fourth sound source vector output by the second gain circuit 31 enters the second band-pass filter 136. The fourth sound source vector has its band limited by the filter 136, whereby a sixth sound source vector is obtained. The second band-pass filter 136 outputs the sixth sound source vector to the adder 40.

[0116] The adder 40 adds the inputs applied thereto, namely the fifth sound source vector output by the first band-pass filter 135 and the sixth sound source vector output by the second band-pass filter 136, and outputs an excitation vector, which is the sum of the fifth and sixth sound source vectors, to the linear prediction filter 150.

[0117] The linear prediction filter 150 has a table in which quantized values of linear prediction coefficients have been stored. The excitation vector output by the adder 40 and an index corresponding to a quantized value of a linear prediction coefficient output by the first linear prediction coefficient calculation circuit 140 enter the linear prediction filter 150. The linear prediction filter 150 reads the quantized value of the linear prediction coefficient corresponding to this index out of the table and drives the filter thus set to this quantized linear prediction coefficient by the excitation vector, whereby a reconstructed vector is obtained. The linear prediction filter 150 outputs this reconstructed vector to the subtractor 50.

[0118] The input vector enters the subtractor 50 via the input terminal 10, and the reconstructed vector output by the linear prediction filter 150 also enters the subtractor 50. The subtractor 50 calculates the difference between these two inputs. The subtractor 50 outputs a difference vector, which is the difference between the input vector and the reconstructed vector, to the weight-

55

ing filter 160.

[0119] The difference vector output by the subtractor 50 and the linear prediction coefficient output by the first linear prediction coefficient calculation circuit 140 enter the weighting filter 160. The latter uses this linear prediction coefficient to produce a weighting filter corresponding to the characteristic of the human sense of hearing and drives this weighting filter by the difference vector, whereby there is obtained a weighted difference vector. The weighted difference vector is output to the minimizing circuit 170.

[0120] The weighted difference vector output by the weighting filter 160 enters the minimizing circuit 170, which proceeds to calculate the norm. Indices corresponding to all values of the elements of the first pulse position vector in the first pulse position generating circuit 110 are output successively from the minimizing circuit 170 to the first pulse position generating circuit 110. Indices corresponding to all values of the elements of the second pulse position vector in the second pulse position generating circuit 111 are output successively from the minimizing circuit 170 to the second pulse position generating circuit 111. Indices corresponding to all first pulse amplitude vectors that have been stored in the first pulse amplitude generating circuit 120 are output successively from the minimizing circuit 170 to the first pulse amplitude generating circuit 120. Indices corresponding to all second pulse amplitude vectors that have been stored in the second pulse amplitude generating circuit 121 are output successively from the minimizing circuit 170 to the second pulse amplitude generating circuit 121. Indices corresponding to all first gains that have been stored in the first gain circuit 30 are output successively from the minimizing circuit 170 to the first gain circuit 30. Indices corresponding to all second gains that have been stored in the second gain circuit 31 are output successively from the minimizing circuit 170 to the second gain circuit 31. Further, the minimizing circuit 170 selects the value of each element in the first pulse position vector, the amount of pulse position revision, the first pulse amplitude vector, the second pulse amplitude vector and the first gain and second gain that will result in the minimum norm and outputs the indices corresponding to these to the code output circuit 190.

[0121] The index corresponding to the quantized value of the linear prediction coefficients output by the first linear prediction coefficient calculation circuit 140 enters the code output circuit 190 and so do the indices corresponding to the value of each element in the first pulse position vector, the amount of pulse position revision, the first pulse amplitude vector, the second pulse amplitude vector and the first gain and second gain. The code output circuit 190 converts each index to a bit-sequence code and outputs the code via the output terminal 60.

[0122] Fig. 2 is a block diagram illustrating the construction of a first embodiment of an apparatus for

encoding speech and musical signals according to the present invention. Components in Fig. 2 identical with or equivalent to those of Fig. 1 are designated by like reference characters.

5 [0123] As shown in Fig. 2, a code in the form of a bit sequence enters from the input terminal 200. A code input circuit 220 converts the bit-sequence code that has entered from the input terminal 200 to an index.

10 [0124] The code input circuit 220 outputs an index corresponding to each element in the first pulse position vector to the first pulse position generating circuit 210; outputs an index corresponding to the amount of pulse position revision to the second pulse position generating circuit 211; outputs an index corresponding to the first pulse amplitude vector to the first pulse amplitude generating circuit 120; outputs an index corresponding to the second pulse amplitude vector to the second pulse amplitude generating circuit 121; outputs an index corresponding to the first gain to the first gain circuit 30; outputs an index corresponding to the second gain to the second gain circuit 31; and outputs an index corresponding to the quantized value of a linear prediction coefficient to the linear prediction filter 150.

15 [0125] The index output by the code input circuit 220 enters the first pulse position generating circuit 210, which proceeds to generate the first pulse position vector using the position of each pulse specified by the index and output the vector to the first sound source generating circuit 20 and to the second pulse position generating circuit 211.

20 [0126] The first pulse amplitude generating circuit 120 has a table in which M-dimensional vectors  $A^-_j$ ,  $j = 1, \dots, N_A$  have been stored. The index output by the code input circuit 220 enters the first pulse amplitude generating circuit 120, which reads an M-dimensional vector  $A^-_j$  corresponding to this index out of the above-mentioned table and outputs this vector to the first sound source generating circuit 20 as a first pulse amplitude vector.

25 [0127] The index output by the code input circuit 220 and the first pulse position vector  $P^- = (P_1, P_2, \dots, P_M)$  output by the first pulse position generating circuit 210 enter the second pulse position generating circuit 211. The latter revises the first pulse position vector using the pulse position revision quantity  $d^-_i = (d_{i1}, d_{i2}, \dots, d_{iM})$  specified by the index and outputs the revised vector to the second sound source generating circuit 21 as a second pulse position vector  $Q^- = (P_1+d_{i1}, P_2+d_{i2}, \dots, P_M+d_{iM})$ .

30 [0128] The second pulse amplitude generating circuit 121 has a table in which M-dimensional vectors  $B^-_j$ ,  $j = 1, \dots, N_B$  have been stored. The index output by the code input circuit 220 enters the second pulse amplitude generating circuit 121, which reads an M-dimensional vector  $B^-_j$  corresponding to this index out of the above-mentioned table and outputs this vector to the second sound source generating circuit 21 as a second pulse amplitude vector.

[0129] The first pulse position vector  $P^* = (P_1, P_2, \dots, P_M)$  output by the first pulse position generating circuit 210 and the first pulse amplitude vector  $A^* = (A_{i1}, A_{i2}, \dots, A_{iM})$  output by the first pulse amplitude generating circuit 120 enter the first sound source generating circuit 20. The first sound source generating circuit 20 outputs an N-dimensional vector for which the values of the  $P_1$ st,  $P_2$ nd, ...,  $P_M$ th elements are  $A_{i1}, A_{i2}, \dots, A_{iM}$ , respectively, and the values of the other elements are zero to the first gain circuit 30 as a first sound source vector.

[0130] A second pulse position vector  $Q^* = (Q_1^*, Q_2^*, \dots, Q_M^*)$  output by the second pulse position generating circuit 211 and a second pulse amplitude vector  $B^* = (B_{i1}, B_{i2}, \dots, B_{iM})$  output by the second pulse amplitude generating circuit 121 enter the second sound source generating circuit 21. The second sound source generating circuit 21 outputs an N-dimensional vector for which the values of the  $Q_1^*$ st,  $Q_2^*$ nd, ...,  $Q_M^*$ th elements are  $B_{i1}, B_{i2}, \dots, B_{iM}$ , respectively, and the values of the other elements are zero to the second gain circuit 31 as a second sound source vector.

[0131] The first gain circuit 30 has a table in which gain values have been stored. The index output by the code input circuit 220 and the first sound source vector output by the first sound source generating circuit 20 enter the first gain circuit 30, which reads a first gain corresponding to the index out of the table, multiplies the first gain by the first sound source vector to thereby generate a third sound source vector, and outputs the generated third sound source vector to the first band-pass filter 135.

[0132] The second gain circuit 31 has a table in which gain values have been stored. The index output by the code input circuit 220 and the second sound source vector output by the second sound source generating circuit 21 enter the second gain circuit 31, which reads a second gain corresponding to the index out of the table, multiplies the second gain by the second sound source vector to thereby generate a fourth sound source vector, and outputs the generated fourth sound source vector to the second band-pass filter 136.

[0133] The third sound source vector output by the first gain circuit 30 enters the first band-pass filter 135. The third sound source vector has its band limited by the filter 135, whereby a fifth sound source vector is obtained. The first band-pass filter 135 outputs the fifth sound source vector to the adder 40.

[0134] The fourth sound source vector output by the second gain circuit 31 enters the second band-pass filter 136. The fourth sound source vector has its band limited by the filter 136, whereby a sixth sound source vector is obtained. The second band-pass filter 136 outputs the sixth sound source vector to the adder 40.

[0135] The adder 40 adds the inputs applied thereto, namely the fifth sound source vector output by the first band-pass filter 135 and the sixth sound source vector output by the second band-pass filter 136, and outputs

an excitation vector, which is the sum of the fifth and sixth sound source vectors, to the linear prediction filter 150.

[0136] The linear prediction filter 150 has a table in which quantized values of linear prediction coefficients have been stored. The excitation vector output by the adder 40 and an index corresponding to a quantized value of a linear prediction coefficient output by the code input circuit 220 enter the linear prediction filter 150. The linear prediction filter 150 reads the quantized value of the linear prediction coefficient corresponding to this index out of the table and drives the filter thus set to this quantized linear prediction coefficient by the excitation vector, whereby a reconstructed vector is obtained. The linear prediction filter 150 outputs this reconstructed vector via the output terminal 201.

#### [Second Embodiment]

[0137] Fig. 3 is a block diagram illustrating the construction of a second embodiment of an apparatus for encoding speech and musical signals according to the present invention. Here also it is assumed for the sake of simplicity that the number of bands is two.

[0138] Components in Fig. 3 identical with or equivalent to those of the prior art illustrated in Fig. 10 are designated by like reference characters and are not described again in order to avoid prolixity.

[0139] As shown in Fig. 3, the first pulse position generating circuit 110 receives as an input an index that is output by the minimizing circuit 170, generates a first pulse position vector using the position of each pulse specified by the index and outputs this vector to the first sound source generating circuit 20 and to the second pulse position generating circuit 111.

[0140] The second pulse position generating circuit 111 receives as inputs the index that is output by the minimizing circuit 170 and the first pulse position vector  $P^* = (P_1, P_2, \dots, P_M)$  output by the first pulse position generating circuit 110, revises the first pulse position vector using the pulse position revision quantity  $d^* = (d_{i1}, d_{i2}, \dots, d_{iM})$  specified by the index and outputs the revised vector to the second sound source generating circuit 21 as a second pulse position vector  $Q^* = (P_1+d_{i1}, P_2+d_{i2}, \dots, P_M+d_{iM})$ .

[0141] The weighted difference vector output by the weighting filter 160 enters the minimizing circuit 170, which proceeds to calculate the norm. Indices corresponding to all values of the elements of the first pulse position vector in the first pulse position generating circuit 110 are output successively from the minimizing circuit 170 to the first pulse position generating circuit 110. Indices corresponding to all values of the elements of the second pulse position vector in the second pulse position generating circuit 111 are output successively from the minimizing circuit 170 to the second pulse position generating circuit 111. Indices corresponding to all first pulse amplitude vectors that have been stored in

the first pulse amplitude generating circuit 120 are output successively from the minimizing circuit 170 to the first pulse amplitude generating circuit 120. Indices corresponding to all second pulse amplitude vectors that have been stored in the second pulse amplitude generating circuit 121 are output successively from the minimizing circuit 170 to the second pulse amplitude generating circuit 121. Indices corresponding to all first gains that have been stored in the first gain circuit 30 are output successively from the minimizing circuit 170 to the first gain circuit 30. Indices corresponding to all second gains that have been stored in the second gain circuit 31 are output successively from the minimizing circuit 170 to the second gain circuit 31. Further, the minimizing circuit 170 selects the value of each element in the first pulse position vector, the amount of pulse position revision, the first pulse amplitude vector, the second pulse amplitude vector and the first gain and second gain that will result in the minimum norm and outputs the indices corresponding to these to the code output circuit 190.

[0142] The index corresponding to the quantized value of the linear prediction coefficient output by the first linear prediction coefficient calculation circuit 140 enters the code output circuit 190 and so do the indices corresponding to the value of each element in the first pulse position vector, the amount of pulse position revision, the first pulse amplitude vector, the second pulse amplitude vector and the first gain and second gain. The code output circuit 190 converts these indices to a bit-sequence code and outputs the code via the output terminal 60.

[0143] Fig. 4 is a block diagram illustrating the construction of the second embodiment of an apparatus for decoding speech and musical signals according to the present invention. Components in Fig. 4 identical with or equivalent to those of Figs. 3 and 12 are designated by like reference characters and are not described again in order to avoid prolixity.

[0144] As shown in Fig. 4, the code input circuit 220 converts the bit-sequence code that has entered from the input terminal 200 to an index. The code input circuit 220 outputs an index corresponding to each element in the first pulse position vector to the first pulse position generating circuit 210, outputs an index corresponding to the amount of pulse position revision to the second pulse position generating circuit 211, outputs an index corresponding to the first pulse amplitude vector to the first pulse amplitude generating circuit 120, outputs an index corresponding to the second pulse amplitude vector to the second pulse amplitude generating circuit 121, outputs an index corresponding to the first gain to the first gain circuit 30, outputs an index corresponding to the second gain to the second gain circuit 31, and outputs an index corresponding to the quantized value of a linear prediction coefficient to the linear prediction filter 150.

[0145] The index output by the code input circuit 220

5 enters the first pulse position generating circuit 210, which generates the first pulse position vector using the position of each pulse specified by the index and outputs the vector to the first sound source generating circuit 20 and to the second pulse position generating circuit 211.

[0146] The index output by the code input circuit 220 and the first pulse position vector  $P^* = (P_1, P_2, \dots, P_M)$  output by the first pulse position generating circuit 210 10 enter the second pulse position generating circuit 211. The latter revises the first pulse position vector using the pulse position revision quantity  $d_i^* = (d_{i1}, d_{i2}, \dots, d_{iM})$  specified by the index and outputs the revised vector to the second sound source generating circuit 21 as a second pulse position vector  $Q^* = (P_1+d_{i1}, P_2+d_{i2}, \dots, P_M+d_{iM})$ .

[Third Embodiment]

20 [0147] Fig. 5 is a block diagram illustrating the construction of a third embodiment of an apparatus for encoding speech and musical signals according to the present invention. As shown in Fig. 5, the apparatus for encoding speech and musical signals according to the third embodiment of the present invention has a higher-order linear prediction coefficient calculation circuit 380 substituted for the higher-order linear prediction coefficient calculation circuit 180 of the second embodiment shown in Fig. 3. Moreover, the first band-pass filter 135 and second band-pass filter 136 are eliminated.

25 [0148] Fig. 6 is a diagram illustrating an example of the construction of the higher-order linear prediction coefficient calculation circuit 380 in the apparatus for encoding speech and musical signals according to the third embodiment depicted in Fig. 5. Components in Fig. 6 identical with or equivalent to those of Fig. 11 are designated by like reference characters and are not described again in order to avoid prolixity. Only the features that distinguish this higher-order linear prediction coefficient calculation circuit will be discussed.

30 [0149] Fourier coefficients output by the FFT circuit 930 enter the band splitting circuit 540. The latter equally partitions these Fourier coefficients into high- and low-frequency regions, thereby obtaining low-frequency Fourier coefficients and high-frequency(region) Fourier coefficients. The low-frequency coefficients are output to the first zerofill circuit 550 and the high-frequency coefficients are output to the second zerofill circuit 551.

35 [0150] The low-frequency Fourier coefficients output by the band splitting circuit 540 enter the first zerofill circuit 550, which fills the band corresponding to the high-frequency region with zeros, generates first full-band Fourier coefficients and outputs these coefficients to the first inverse FFT circuit 560.

40 [0151] The high-frequency Fourier coefficients output by the band splitting circuit 540 enter the second zerofill circuit 551, which fills the band corresponding to the

low-frequency region with zeros, generates second full-band Fourier coefficients and outputs these coefficients to the second inverse FFT circuit 561.

[0152] The first full-band Fourier coefficients output by the first zerofill circuit 550 enter the first inverse FFT circuit 560, which proceeds to subject these coefficients to an inverse FFT, thereby obtaining a first residual signal that is output to the first higher-order linear prediction coefficient calculation circuit 570.

[0153] The second full-band Fourier coefficients output by the second zerofill circuit 551 enter the second inverse FFT circuit 561, which proceeds to subject these coefficients to an inverse FFT, thereby obtaining a second residual signal that is output to the second higher-order linear prediction coefficient calculation circuit 571.

[0154] The first residual signal output by the first inverse FFT circuit 560 enters the first higher-order linear prediction coefficient calculation circuit 570, which proceeds to subject the first residual signal to higher-order linear prediction analysis, thereby obtaining the first higher-order linear prediction coefficient. This is output to the first higher-order linear prediction filter 130 via the output terminal 901.

[0155] The second residual signal output by the second inverse FFT circuit 561 enters the second higher-order linear prediction coefficient calculation circuit 571, which proceeds to subject the second residual signal to higher-order linear prediction analysis, thereby obtaining the second higher-order linear prediction coefficient. This is output to the second higher-order linear prediction filter 131 via the output terminal 902.

[0156] Fig. 7 is a block diagram illustrating the construction of the third embodiment of an apparatus for decoding speech and musical signals according to the present invention. As shown in Fig. 7, the apparatus for decoding speech and musical signals according to the third embodiment of the present invention has the higher-order linear prediction coefficient calculation circuit 380 substituted for the higher-order linear prediction coefficient calculation circuit 180 of the second embodiment shown in Fig. 4.

[0157] Moreover, the first band-pass filter 135 and second band-pass filter 136 are eliminated.

#### [Fourth Embodiment]

[0158] Fig. 8 is a block diagram illustrating the construction of a fourth embodiment of an apparatus for encoding speech and musical signals according to the present invention. As shown in Fig. 8, the apparatus for encoding speech and musical signals according to the fourth embodiment of the present invention has the higher-order linear prediction coefficient calculation circuit 380 substituted for the higher-order linear prediction coefficient calculation circuit 180 shown in Fig. 10. Moreover, the first band-pass filter 135 and second band-pass filter 136 are eliminated.

[0159] Fig. 9 is a block diagram illustrating the construction of the fourth embodiment of an apparatus for decoding speech and musical signals according to the present invention. As shown in Fig. 9, the apparatus for decoding speech and musical signals according to the fourth embodiment of the present invention has the higher-order linear prediction coefficient calculation circuit 380 substituted for the higher-order linear prediction coefficient calculation circuit 180 shown in Fig. 12. Moreover, the first band-pass filter 135 and second band-pass filter 136 are eliminated.

[0160] Though the number of bands is limited to two in the foregoing description for the sake of simplicity, the present invention is applicable in similar fashion to cases where the number of bands is three or more.

[0161] Further, it goes without saying that the present invention may be so adapted that the first pulse position vector is used as the second pulse position vector. Further, it is possible to use all or part of the first pulse amplitude vector as the second pulse amplitude vector.

[0162] Thus, in accordance with the present invention, as described above, the sound source signal of each of a plurality of bands can be encoded using a small number of bits in a band-splitting-type apparatus for encoding speech and musical signals. The reason for this is that the correlation between bands possessed by the input signal is taken into consideration some of the information possessed by a sound source signal that has been encoded in a certain band or bands is used to encode a sound source signal in the other band(s).

[0163] As many apparently widely different embodiments of the present invention can be made without departing from the spirit and scope thereof, it is to be understood that the invention is not limited to the specific embodiments thereof except as defined in the appended claims.

#### Claims

1. A speech and musical signal encoding apparatus which, when encoding an input signal upon splitting the input signal into a plurality of bands, generates a reconstructed signal using a multi pulse sound source signal that corresponds to each band, wherein a position obtained by shifting the position of each pulse which defines the multipulse signal in the band(s) is used when defining a multipulse signal in the other band(s).
2. A speech and musical signal decoding apparatus for generating a reconstructed signal using a multipulse sound source signal corresponding to each of a plurality of bands, wherein a position obtained by shifting the position of each pulse which defines the multipulse signal in the band(s) is used when defining a multipulse signal in the other band(s).

3. A speech and musical signal encoding apparatus which, when encoding an input signal upon splitting the input signal into a plurality of bands, generates a reconstructed signal by exciting a synthesis filter by a full-band sound source signal, which is obtained by summing, over all bands, multipulse sound source signals corresponding to respective ones of the plurality of bands,

wherein a position obtained by shifting the position of each pulse which defines the multipulse signal in the band(s) is used when defining a multipulse signal in the other band(s).

4. A speech and musical signal decoding apparatus for generating a reconstructed signal by exciting a synthesis filter by a full-band sound source signal, which is obtained by summing, over all bands, multipulse sound source signals corresponding to respective ones of a plurality of bands,

wherein a position obtained by shifting the position of each pulse which defines the multipulse signal in the band(s) is used when defining a multipulse signal in the other band(s).

5. A speech and musical signal encoding apparatus which, when encoding an input signal upon splitting the input signal into a plurality of bands, generates a reconstructed signal by exciting a synthesis filter by a full-band sound source signal, which is obtained by summing, over all bands, signals obtained by exciting a higher-order linear prediction filter, which represents a microspectrum relating to the input signal of each band, by a multi pulse sound source signal corresponding to each band,

wherein a position obtained by shifting the position of each pulse which defines the multipulse signal in the band(s) is used when defining a multipulse signal in the other band(s).

6. A speech and musical signal decoding apparatus for generating a reconstructed signal by exciting a synthesis filter by a full-band sound source signal, which is obtained by summing, over all bands, signals obtained by exciting a higher-order linear prediction filter, which represents a microspectrum relating to an input signal of each of a plurality of bands, by a multi pulse sound source signal corresponding to each band,

wherein a position obtained by shifting the position of each pulse which defines the multipulse signal in the band(s) is used when defining a multipulse signal in the other band(s).

7. A speech and musical signal encoding apparatus which, when encoding an input signal upon splitting the input signal into a plurality of bands, generates a reconstructed signal by exciting a synthesis filter by a full-band sound source signal, which is

5

obtained by summing, over all bands, signals obtained by exciting a higher-order linear prediction filter, which represents a microspectrum relating to the input signal of each band, by a multipulse sound source signal corresponding to each band, wherein:

a residual signal is found by inverse filtering of the reconstructed signal using a linear prediction filter for which linear prediction coefficients obtained from the reconstructed signal have been decided; and

conversion coefficients obtained by converting the residual signal are split into bands, and said higher-order linear prediction filter uses coefficients obtained from a residual signal of each band generated in each band by back-converting the conversion coefficients that have been split into the bands.

10

15

20

25

30

35

40

45

50

55

8. A speech and musical signal decoding apparatus for generating a reconstructed signal by exciting a synthesis filter by a full-band sound source signal, which is obtained by summing, over all bands, signals obtained by exciting a higher-order linear prediction filter, which represents a microspectrum relating to an input signal of each of a plurality of bands, by a multipulse sound source signal corresponding to each band, wherein:

a residual signal is found by inverse filtering of the reconstructed signal using a linear prediction filter for which linear prediction coefficients obtained from the reconstructed signal have been decided; and

conversion coefficients obtained by converting the residual signal are split into bands, and said higher-order linear prediction filter uses coefficients obtained from a residual signal of each band generated in each band by back-converting the conversion coefficients that have been split into the bands.

35

40

45

50

55

9. The apparatus according to claim 5 or 6, wherein a residual signal is found by inverse filtering of the reconstructed signal using a linear prediction filter for which linear prediction coefficients obtained from the reconstructed signal have been decided,

conversion coefficients obtained by converting the residual signal are split into bands, and said higher-order linear prediction filter uses coefficients obtained from a residual signal of each band generated in each band by back-converting the conversion coefficients that have been split into the bands.

10. A speech and musical signal encoding apparatus

which, when encoding an input signal upon splitting the input signal into a plurality of bands, generates a reconstructed signal using a multipulse sound source signal that corresponds to each band, comprising:

5

(a) first pulse position generating means, to which an index output by minimizing means is input, for generating a first pulse position vector using the position of each pulse specified by the index and outputting the first pulse position vector to a corresponding sound source generating means and to one or a plurality of other pulse position generating means; and  
 (b) one or a plurality of pulse position generating means, to which the index output by said minimizing means and the first pulse position vector output by said first pulse position generating means are input, for generating a pulse position vector by revising the first pulse position vector using a pulse position revision quantity specified by the index, and outputting this revised pulse position vector to corresponding sound source generating means.

10

(b) one or a plurality of pulse position generating means, to which the index output by said minimizing means and the first pulse position vector output by said first pulse position generating means are input, for generating a pulse position vector by revising the first pulse position vector using a pulse position revision quantity specified by the index, and outputting this revised pulse position vector to corresponding sound source generating means.

15

(c) first and second pulse amplitude generating means, to which the index output by said minimizing means is input, for outputting first and second pulse amplitude vectors to said first and second sound source generating means, respectively, from said index;

20

(d) said first and second sound source generating means, to which the first and second pulse position vectors output by said first and second pulse position generating means and the first and second pulse amplitude vectors output by said first and second pulse amplitude generating means are respectively input, for generating first and second sound source vectors and outputting the first and second sound source vectors to first and second gain means, respectively;

25

(e) first and second gain means, each of which has a table in which gain values have been stored and to which the index output by said minimizing means and the first and second sound source vectors, respectively, output by said first and second sound source generating means are input, for reading first and second gains corresponding to the index out of the tables, multiplying the first and second gains by the first and second sound source vectors, respectively, and outputting the products as third and fourth sound source vectors, respectively;

30

(f) first and second band-pass filters for band-passing the third and fourth sound source vectors from said first and second gain means and outputting them as fifth and sixth sound source vectors, respectively;

35

(g) adding means for adding the fifth and sixth sound source vectors output thereto from said first and second band-pass filters, respectively, and outputting an excitation vector, which is the sum of the fifth and sixth sound source vectors, to a linear prediction filter;

40

(h) a linear prediction filter, which has a table in which quantized values of linear prediction coefficients have been stored and to which the excitation vector output by said adding means and an index corresponding to a quantized value of a linear prediction coefficient output by

11. A speech and musical signal decoding apparatus for generating a reconstructed signal using a multipulse sound source signal corresponding to each of a plurality of bands, comprising:

45

(a) first pulse position generating means, to which an index output by code input means is input, for generating a first pulse position vector using the position of each pulse specified by the index and outputting the first pulse position vector to a corresponding sound source generating means and to one or a plurality of other pulse position generating means; and  
 (b) one or a plurality of pulse position generating means, to which the index output by said code input means and the first pulse position vector output by said first pulse position generating means are input, for generating a pulse position vector by revising the first pulse position vector using a pulse position revision quantity specified by the index, and outputting this revised pulse position vector to corresponding sound source generating means.

50

(c) first and second pulse amplitude generating means, to which the index output by said minimizing means is input, for outputting first and second pulse amplitude vectors to said first and second sound source generating means, respectively, from said index;

55

(d) said first and second sound source generating means, to which the first and second pulse position vectors output by said first and second pulse position generating means and the first and second pulse amplitude vectors output by said first and second pulse amplitude generating means are respectively input, for generating first and second sound source vectors and outputting the first and second sound source vectors to first and second gain means, respectively;

(e) first and second gain means, each of which has a table in which gain values have been stored and to which the index output by said minimizing means and the first and second sound source vectors, respectively, output by said first and second sound source generating means are input, for reading first and second gains corresponding to the index out of the tables, multiplying the first and second gains by the first and second sound source vectors, respectively, and outputting the products as third and fourth sound source vectors, respectively;

(f) first and second band-pass filters for band-passing the third and fourth sound source vectors from said first and second gain means and outputting them as fifth and sixth sound source vectors, respectively;

(g) adding means for adding the fifth and sixth sound source vectors output thereto from said first and second band-pass filters, respectively, and outputting an excitation vector, which is the sum of the fifth and sixth sound source vectors, to a linear prediction filter;

(h) a linear prediction filter, which has a table in which quantized values of linear prediction coefficients have been stored and to which the excitation vector output by said adding means and an index corresponding to a quantized value of a linear prediction coefficient output by

12. A speech and music encoding apparatus comprising:

50

(a) first pulse position generating means, to which an index output by minimizing means is input, for generating a first pulse position vector using the position of each pulse specified by the index and outputting the first pulse position vector to first sound source generating means

55

and to second pulse position generating means;

(b) second pulse position generating means, to which the index output by said minimizing means and the first pulse position vector output by said first pulse position generating means are input, for revising the first pulse position vector using a pulse position revision quantity specified by the index, and outputting this revised pulse position vector to second sound source generating means as a second pulse position vector;

(c) first and second pulse amplitude generating means, to which the index output by said minimizing means is input, for outputting first and second pulse amplitude vectors to said first and second sound source generating means, respectively, from said index;

(d) said first and second sound source generating means, to which the first and second pulse position vectors output by said first and second pulse position generating means and the first and second pulse amplitude vectors output by said first and second pulse amplitude generating means are respectively input, for generating first and second sound source vectors and outputting the first and second sound source vectors to first and second gain means, respectively;

(e) first and second gain means, each of which has a table in which gain values have been stored and to which the index output by said minimizing means and the first and second sound source vectors, respectively, output by said first and second sound source generating means are input, for reading first and second gains corresponding to the index out of the tables, multiplying the first and second gains by the first and second sound source vectors, respectively, and outputting the products as third and fourth sound source vectors, respectively;

(f) first and second band-pass filters for band-passing the third and fourth sound source vectors from said first and second gain means and outputting them as fifth and sixth sound source vectors, respectively;

(g) adding means for adding the fifth and sixth sound source vectors output thereto from said first and second band-pass filters, respectively, and outputting an excitation vector, which is the sum of the fifth and sixth sound source vectors, to a linear prediction filter;

(h) a linear prediction filter, which has a table in which quantized values of linear prediction coefficients have been stored and to which the excitation vector output by said adding means and an index corresponding to a quantized value of a linear prediction coefficient output by

first linear prediction coefficient calculation means are input, for reading a quantized value of a linear prediction coefficient corresponding to said index out of the table and driving a filter, for which this quantized linear prediction coefficient has been set, by the excitation vector, thereby obtaining a reconstructed vector, said reconstructed vector being output to subtraction means;

(i) first linear prediction coefficient calculation means for obtaining a linear prediction coefficient by applying linear prediction analysis to an input vector from an input terminal, quantizing this linear prediction coefficient, outputting this linear prediction coefficient to a weighting filter and outputting an index, which corresponds to the quantized value of this linear prediction coefficient, to a linear prediction filter and to code output means;

(j) subtraction means, to which an input vector is input via the input terminal and to which the reconstructed vector output by said linear prediction filter is input, for outputting a difference vector, which is the difference between the input vector and the reconstructed vector, to the weighting filter;

(k) said weighting filter, to which the difference vector output by said difference means and the linear prediction coefficient output by said first linear prediction calculating means are input, for generating a weighting filter corresponding to the characteristic of the human sense of hearing using this linear prediction coefficient and driving said weighting filter by the difference vector, thereby obtaining a weighted difference vector, said weighted difference vector being output to said minimizing means;

(l) minimizing means, to which weighted difference vectors output by said weighting filter are successively input, for calculating norms of these vectors; successively outputting, to said first pulse position generating means, indices corresponding to all values of the elements in the first pulse position vector; successively outputting, to said second pulse position generating means, indices corresponding to all pulse position revision quantities; successively outputting, to said first pulse amplitude generating means, indices corresponding to all first pulse amplitude vectors; successively outputting, to said second pulse amplitude generating means, indices corresponding to all second pulse amplitude vectors;

successively outputting, to said first gain means, indices corresponding to all first gains; successively outputting, to said second gain means, indices corresponding to all second gains; selecting, so as to minimize the norms,

5

10

15

20

25

30

35

40

45

50

55

the value of each element in the first pulse position vector, the pulse position revision quantity, the first pulse amplitude vector, the second pulse amplitude vector and the first gain and second gain; and outputting indices corresponding to these to said code output means; and

(m) code output means, to which the index corresponding to the quantized value of the linear prediction coefficient output by said first linear prediction coefficient calculation means is input as well as the indices, which are output by said minimizing means, corresponding to the value of each element in the first pulse position vector, the pulse position revision quantity, the first pulse amplitude vector, the second pulse amplitude vector and the first gain and second gain, respectively, for converting each index to a bit-sequence code and outputting the bit-sequence code from an output terminal.

13. A speech and music decoding apparatus comprising:

(a) code input means for converting a bit-sequence code, which has entered from an input terminal, to an index;

(b) first pulse position generating means, to which an index output by said code input means is input, for generating a first pulse position vector using the position of each pulse specified by the index and outputting the first pulse position vector to first sound source generating means and to second pulse position generating means;

(c) second pulse position generating means, to which the index output by said code input means and the first pulse position vector output by said first pulse position generating means are input, for revising the first pulse position vector using a pulse position revision quantity specified by the index, and outputting this revised pulse position vector to second sound source generating means as a second pulse position vector;

(d) first and second pulse amplitude generating means, to which the index output by said code input means is input, for reading out vectors corresponding to this index and outputting these vectors to first and second pulse amplitude generating means as first and second amplitude vectors, respectively;

(e) first and second sound source generating means, to which the first and second pulse position vectors output by said first and second pulse position generating means and the first and second pulse amplitude vectors output by said first and second pulse amplitude generat-

ing means are respectively input, for generating first and second sound source vectors and outputting the first and second sound source vectors to first and second gain means, respectively;

(f) first and second gain means, each of which has a table in which gain values have been stored and to which the index output by said code input means and the first and second sound source vectors, respectively, output by said first and second sound source generating are input, for reading first and second gains corresponding to the index out of the tables, multiplying the first and second gains by the first and second sound source vectors, respectively, to thereby generate third and fourth sound source vectors, and outputting the generated third and fourth sound source vectors to first and second band-pass filters, respectively;

(g) adding means for adding the fifth and sixth sound source vectors output thereto from said first and second band-pass filters, respectively, and outputting an excitation vector, which is the sum of the fifth and sixth sound source vectors, to a linear prediction filter; and

(h) a linear prediction filter, which has a table in which quantized values of linear prediction coefficients have been stored and to which the excitation vector output by said adding means and an index corresponding to a quantized value of a linear prediction coefficient output by first linear prediction coefficient calculation means are input, for reading a quantized value of a linear prediction coefficient corresponding to said index out of the table and driving a filter, for which this quantized linear prediction coefficient has been set, by the excitation vector, thereby obtaining a reconstructed vector, said reconstructed vector being output from an output terminal.

14. The apparatus according to claim 12 or 13, further comprising first and second higher-order linear prediction filters to which the third and fourth sound source vectors respectively generated by said first and second gain means are input, respectively;

wherein third and fourth higher-order linear prediction coefficients output from higher-order linear prediction coefficient calculating means whose input is the output of said linear prediction filter, as well as the third and fourth sound source vectors respectively output by said first and second gains means, are respectively input to said first and second higher-order linear prediction filters, said first and second higher-order linear prediction filters driving filters, for which the third and fourth higher-order linear prediction coefficients have been set, by the third and fourth sound source vectors,

respectively, thereby to obtain first and second excitation vectors that are output to said first and second band-pass filters, respectively.

5 15. The apparatus according to claim 12 or 13, wherein said first and second band-pass filters are deleted, and outputs of said first and second higher-order linear prediction filters are input to said adding means.

10 16. The apparatus according to claim 12 or 13, further comprising:

15 second linear prediction coefficient calculation means, to which the reconstructed vector output by said linear prediction filter is input, for applying linear prediction analysis to the reconstructed vector and obtaining a second linear prediction coefficient;

20 residual signal calculation means, to which the second linear prediction coefficient output by said second linear prediction coefficient calculation means and the reconstructed vector output by said linear prediction filter are input, for outputting a residual vector by subjecting the reconstructed vector to inverse filtering processing using a filter for which the second linear prediction coefficient has been set;

25 FFT means, to which the residual vector from said residual signal calculation means is input, for subjecting the residual vector to a fast-Fourier transform;

30 band splitting means, to which Fourier coefficients output by said FFT means are input, for equally partitioning these Fourier coefficients into low- and high-frequency regions to obtain low-frequency Fourier coefficients and high-frequency Fourier coefficients, and for outputting these low-frequency Fourier coefficients and high-frequency Fourier coefficients;

35 first zerofill means, to which the low-frequency Fourier coefficients output by said band splitting means are input, for filling the band corresponding to the high-frequency region with zeros to thereby generate and output first full-band Fourier coefficients;

40 second zerofill means, to which the high-frequency Fourier coefficients output by said band splitting means are input, for filling the band corresponding to the low-frequency region with zeros to thereby generate and output second full-band Fourier coefficients;

45 first inverse FFT means, to which the first full-band Fourier coefficients output by said first zerofill means are input, for subjecting these coefficients to an inverse fast-Fourier transform and outputting a first residual signal thus obtained;

second inverse FFT means, to which the second full-band Fourier coefficients output by said second zerofill means are input, for subjecting these coefficients to an inverse fast-Fourier transform and outputting a second residual signal thus obtained; 5

first higher-order linear prediction coefficient calculation means, to which the first residual signal is input, for applying higher-order linear prediction analysis to the first residual signal to obtain a first higher-order linear prediction coefficient, and outputting this coefficient to said first higher-order linear prediction filter; and 10

second higher-order linear prediction coefficient calculation means, to which the second residual signal is input, for applying higher-order linear prediction analysis to the second residual signal to obtain a second higher-order linear prediction coefficient, and outputting this coefficient to said second higher-order linear prediction filter. 15 20

25

30

35

40

45

50

55

FIG. 1

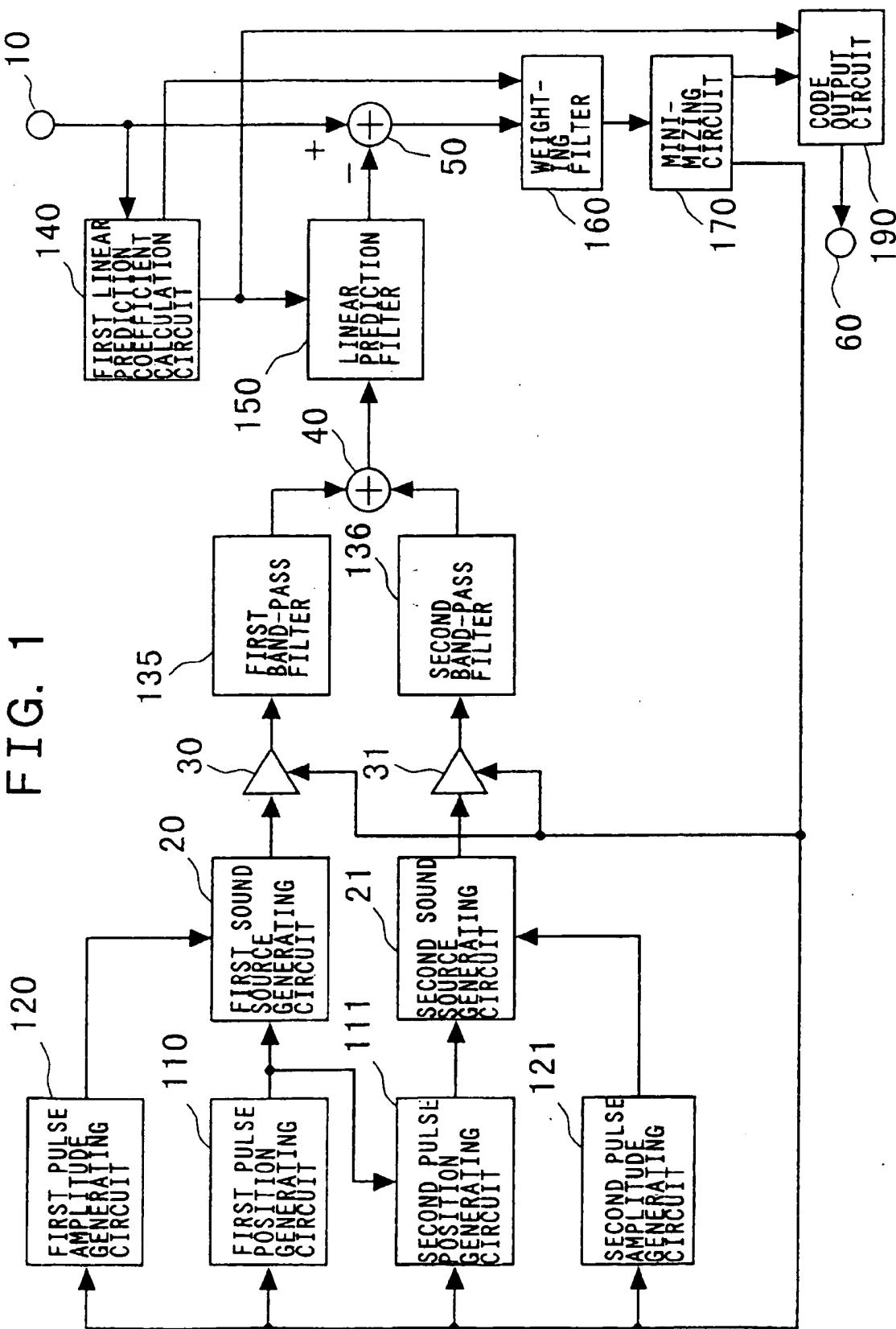


FIG. 2

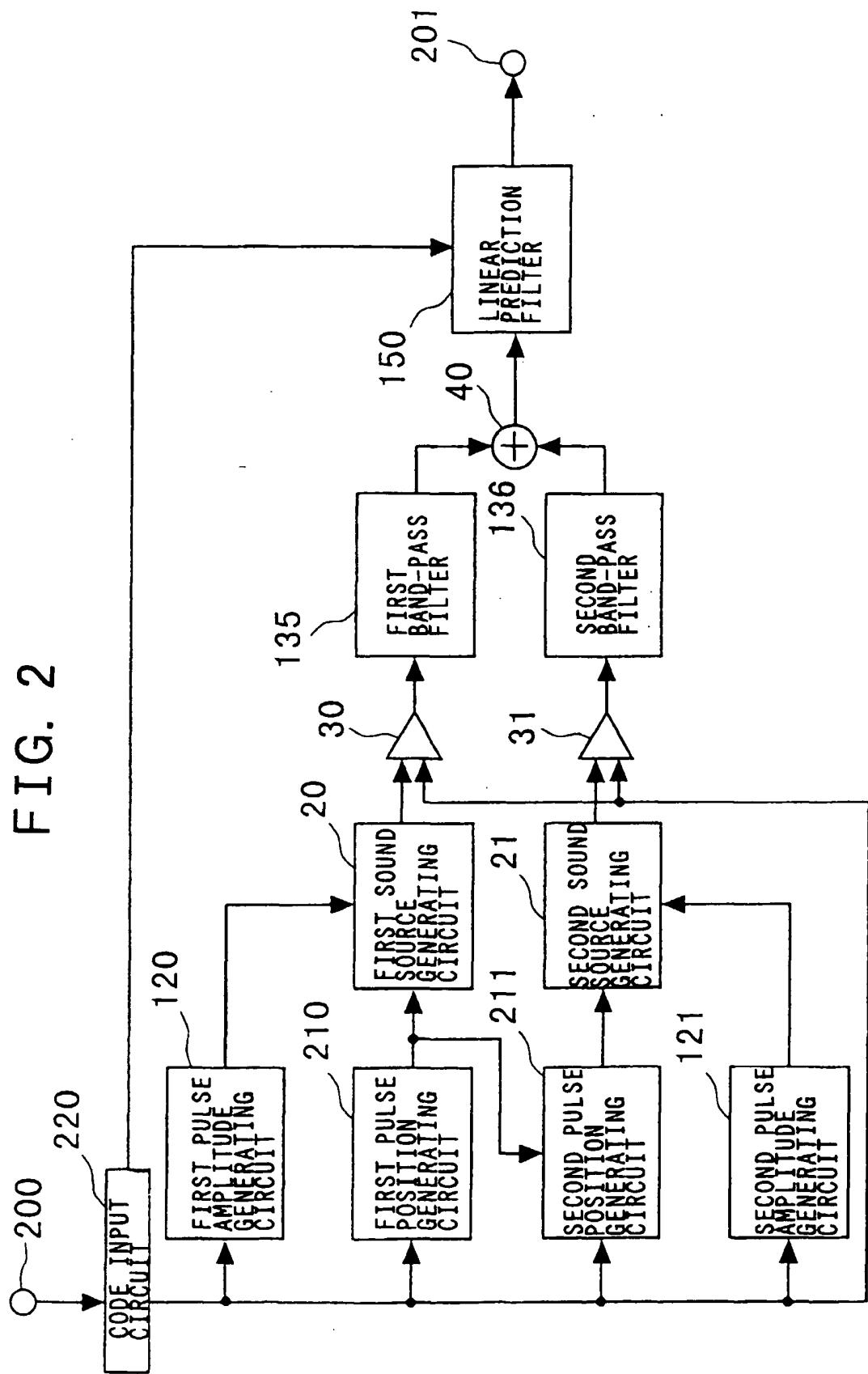


FIG. 3

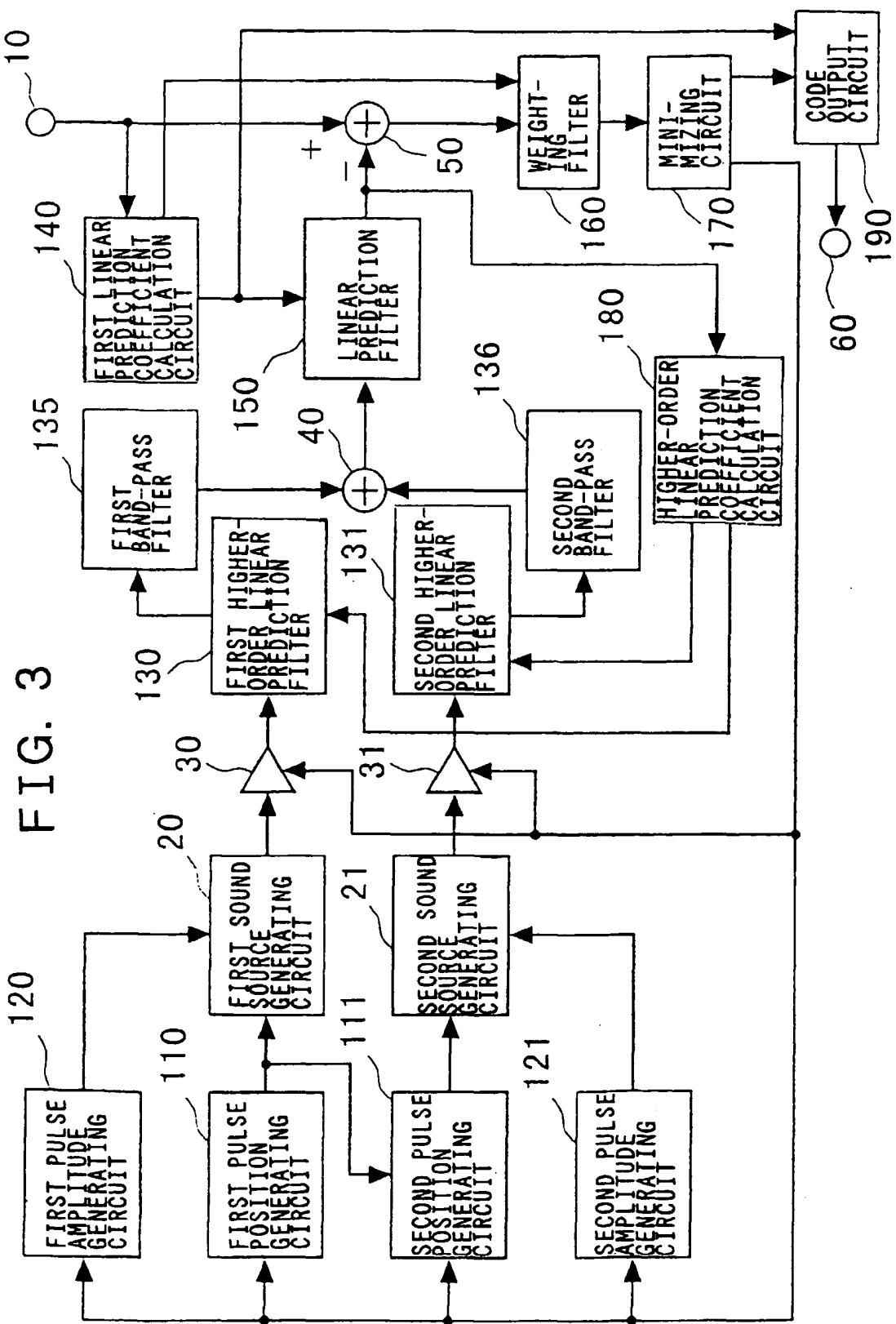


FIG. 4

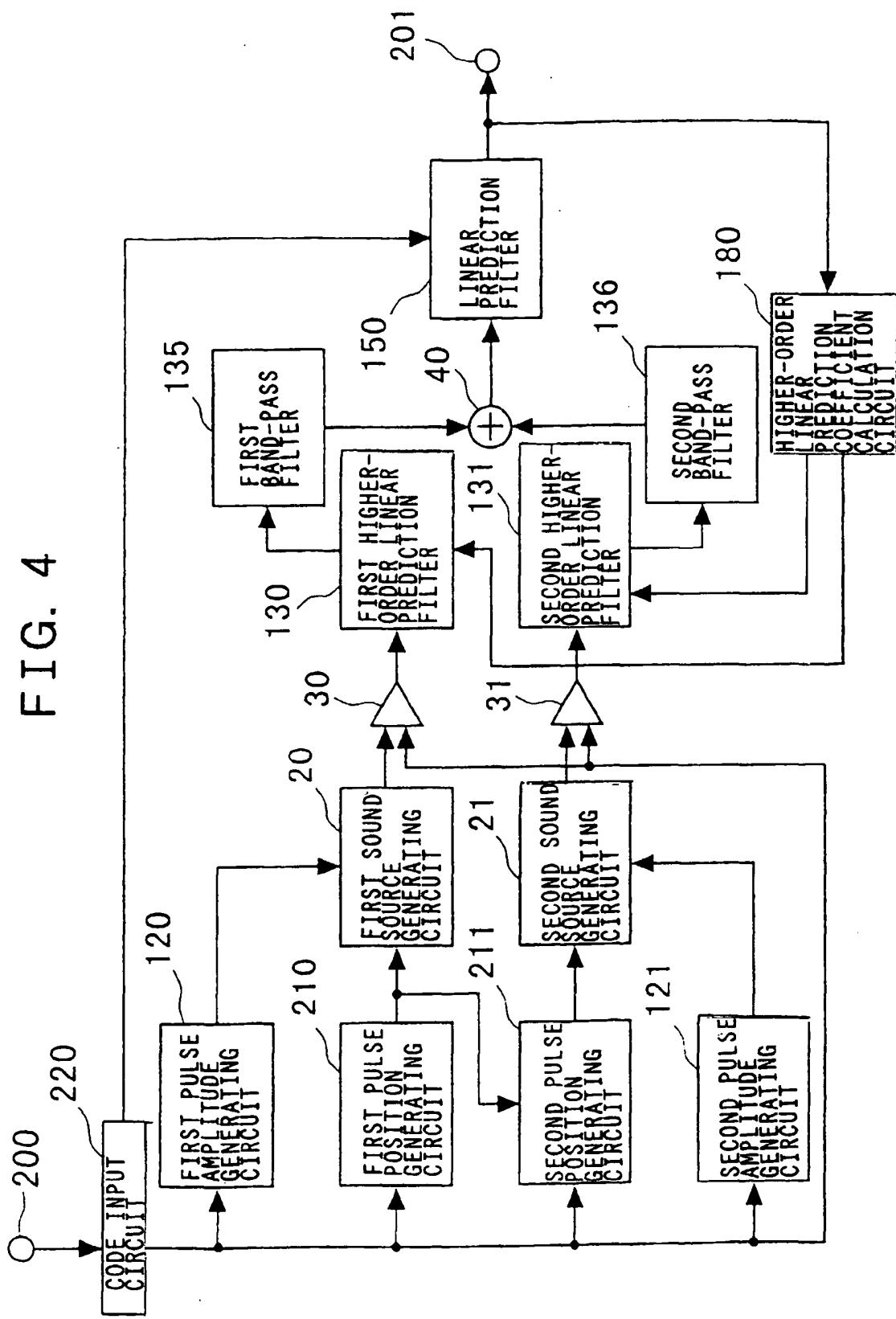


FIG. 5

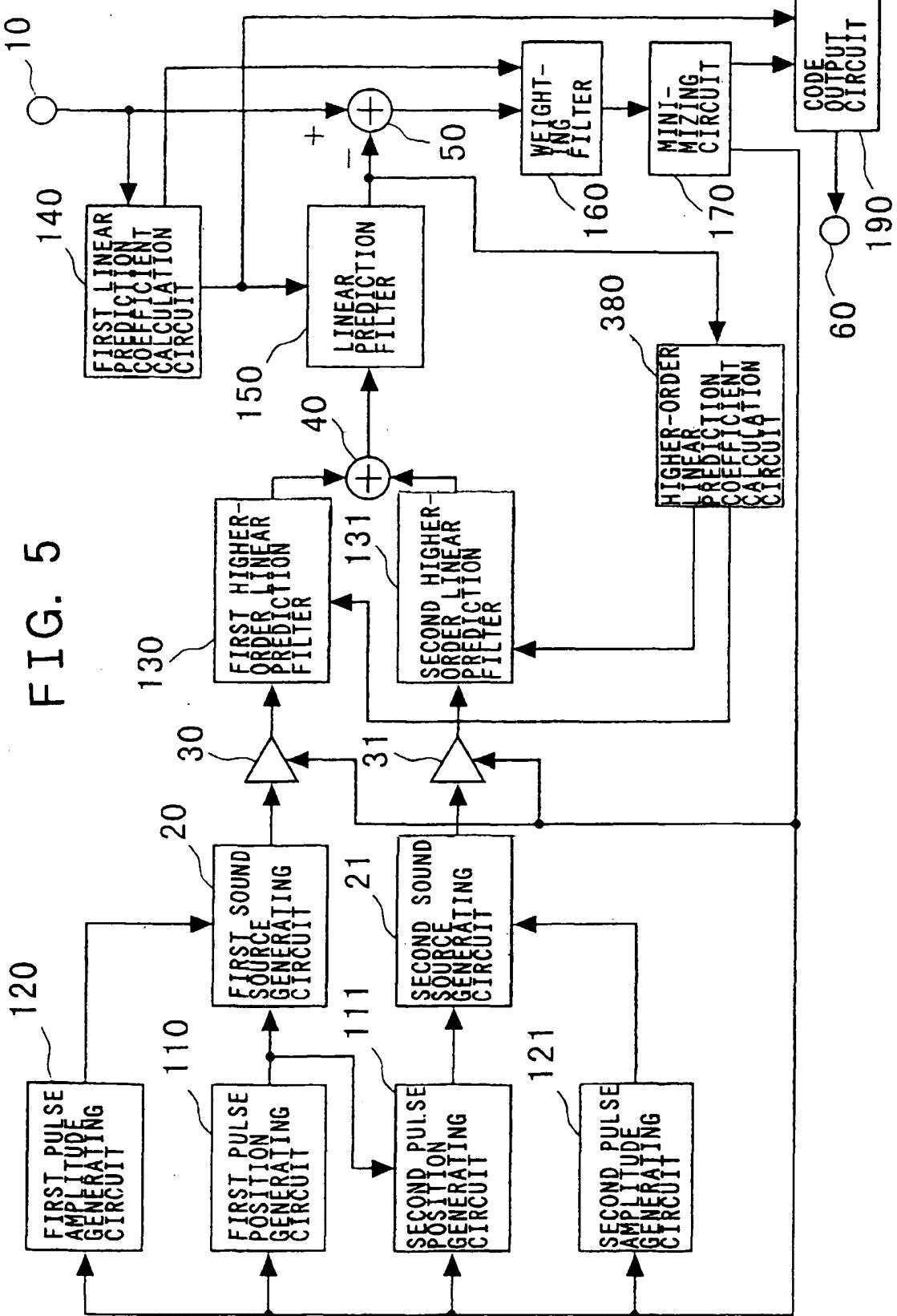


FIG. 6

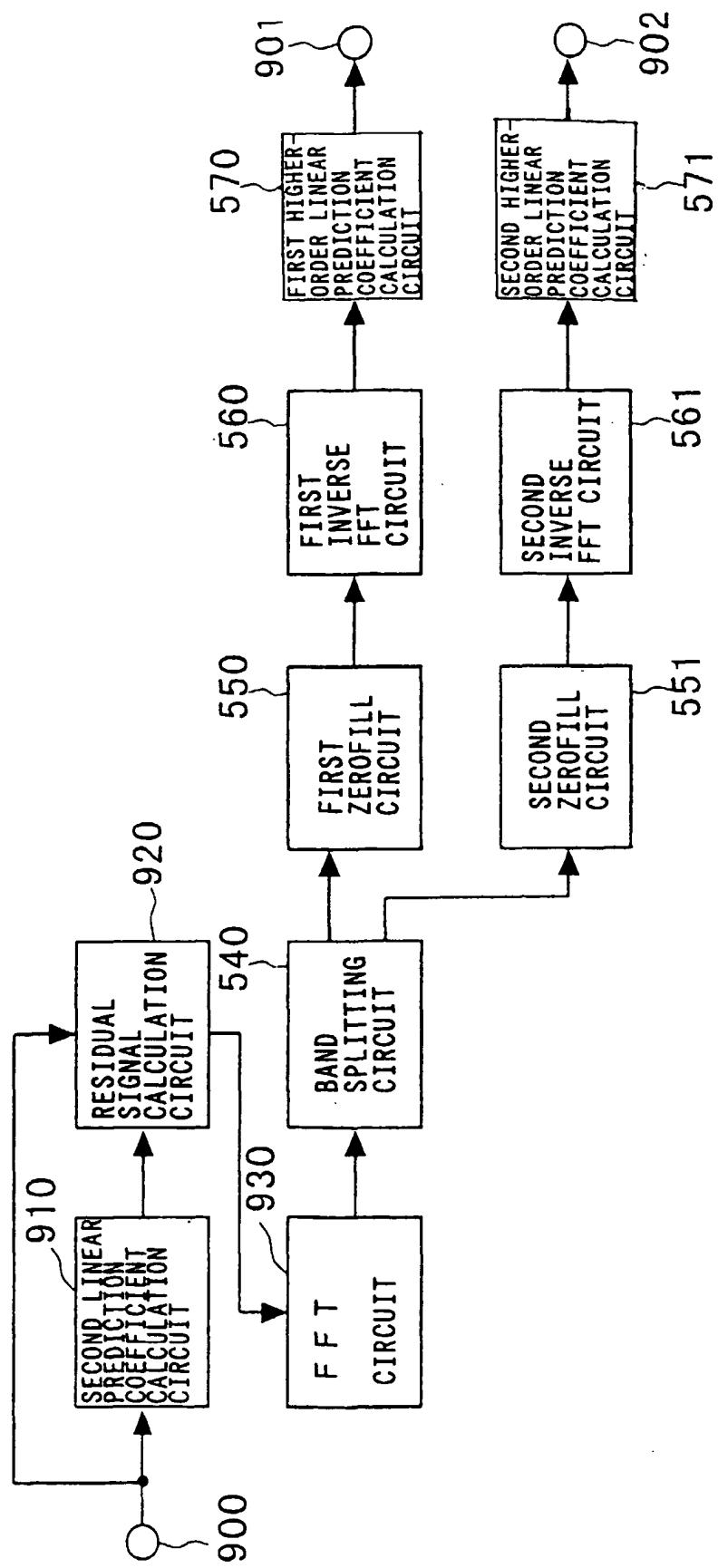


FIG. 7

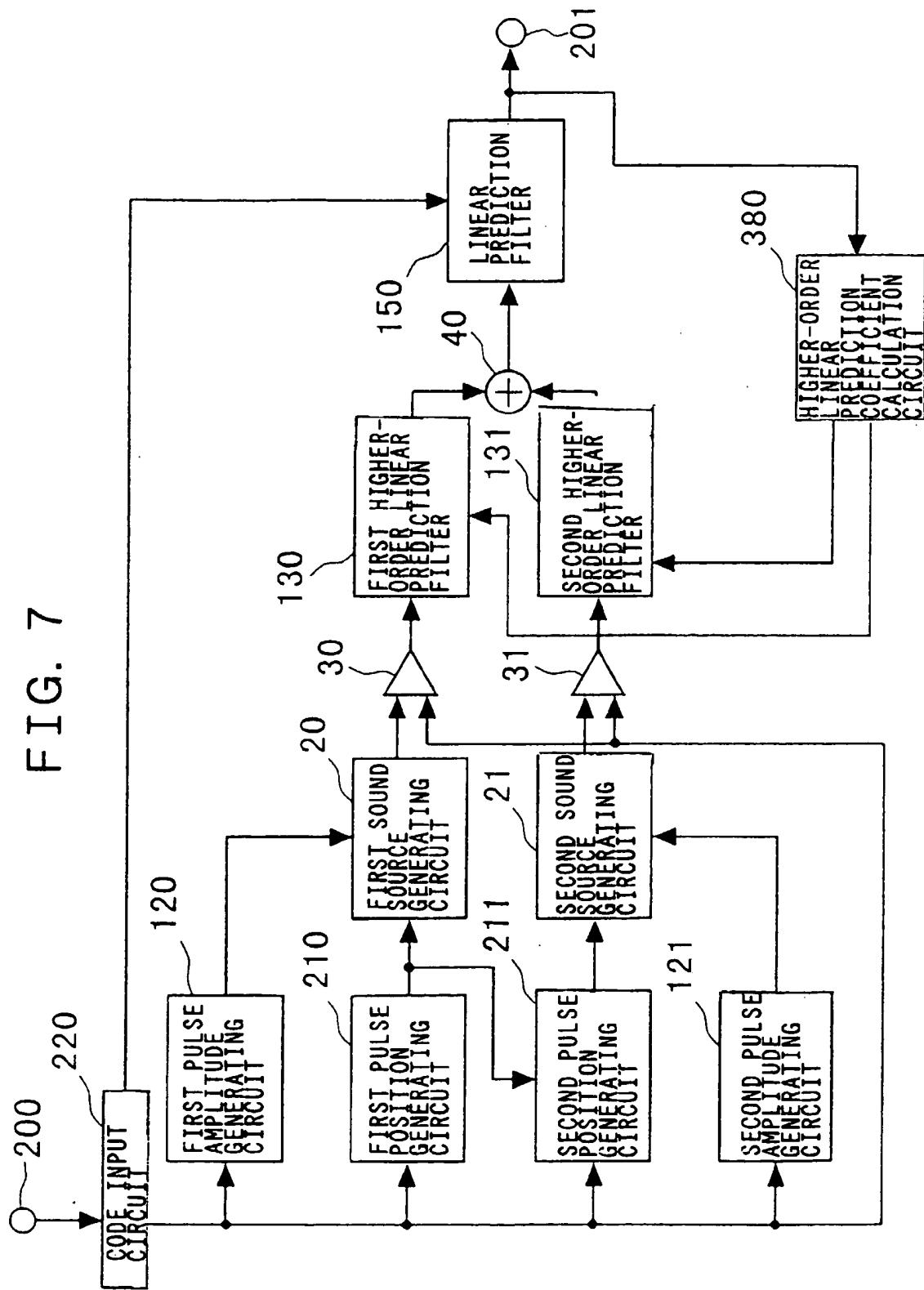


FIG. 8

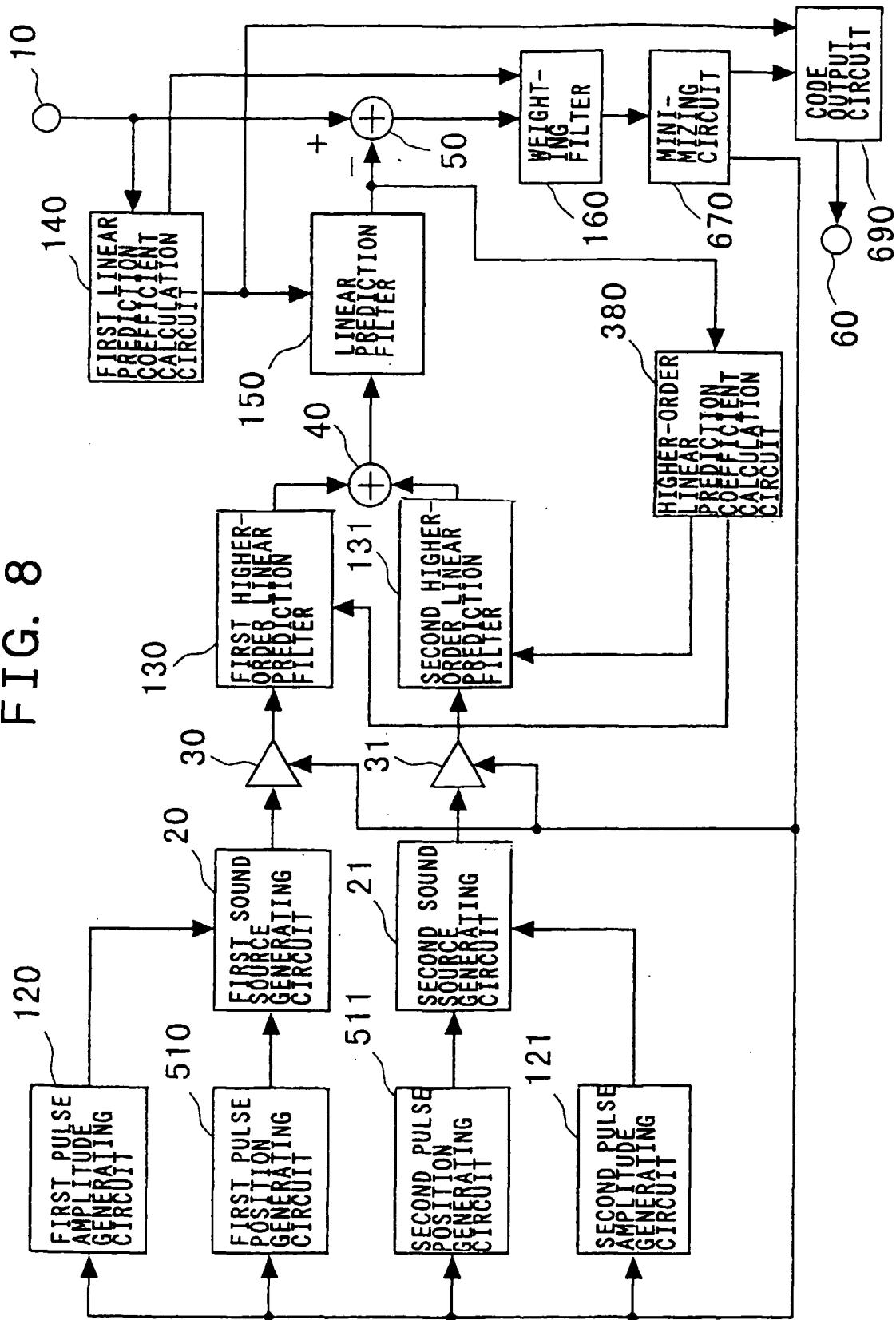


FIG. 9

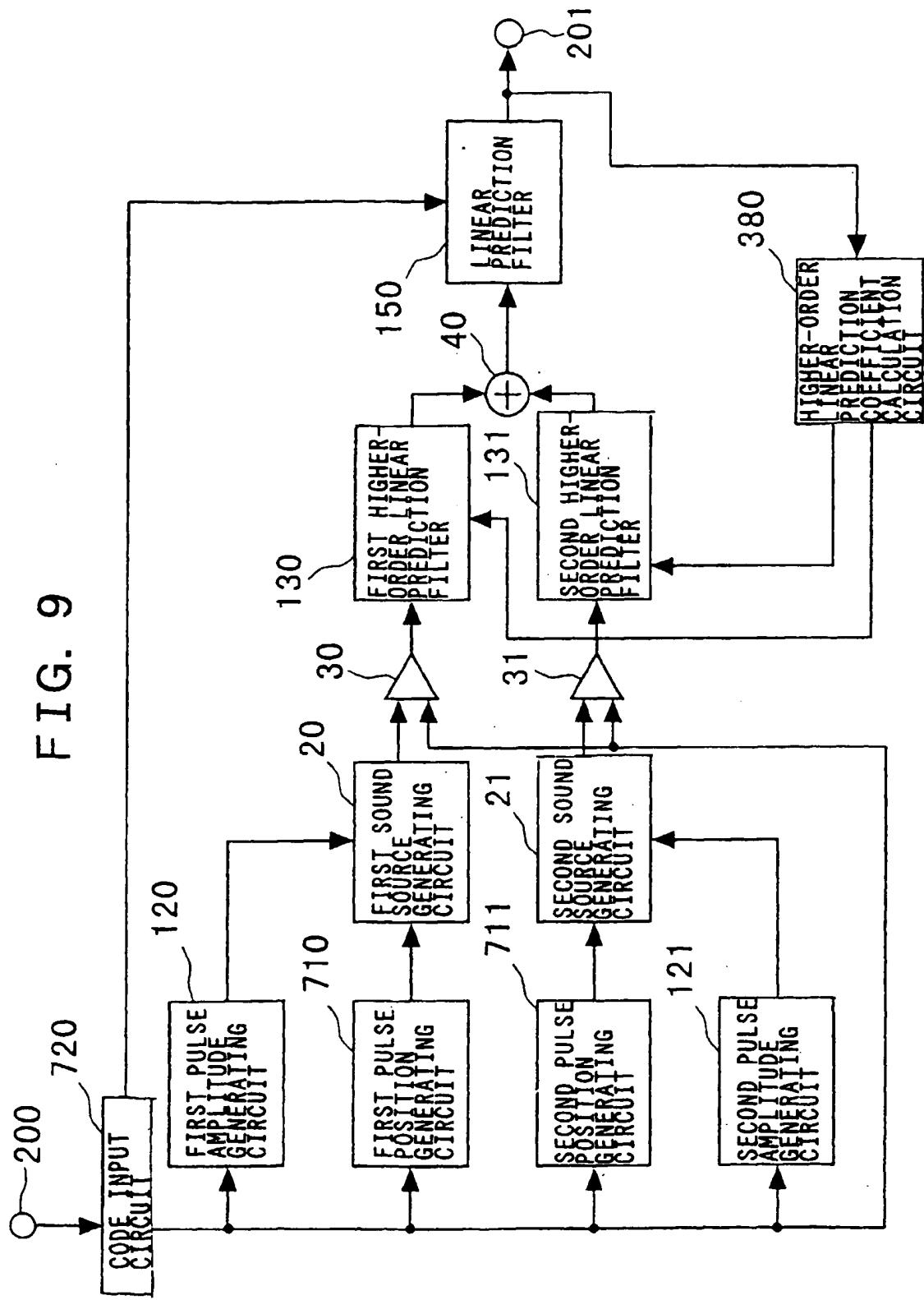


FIG. 10 PRIOR ART

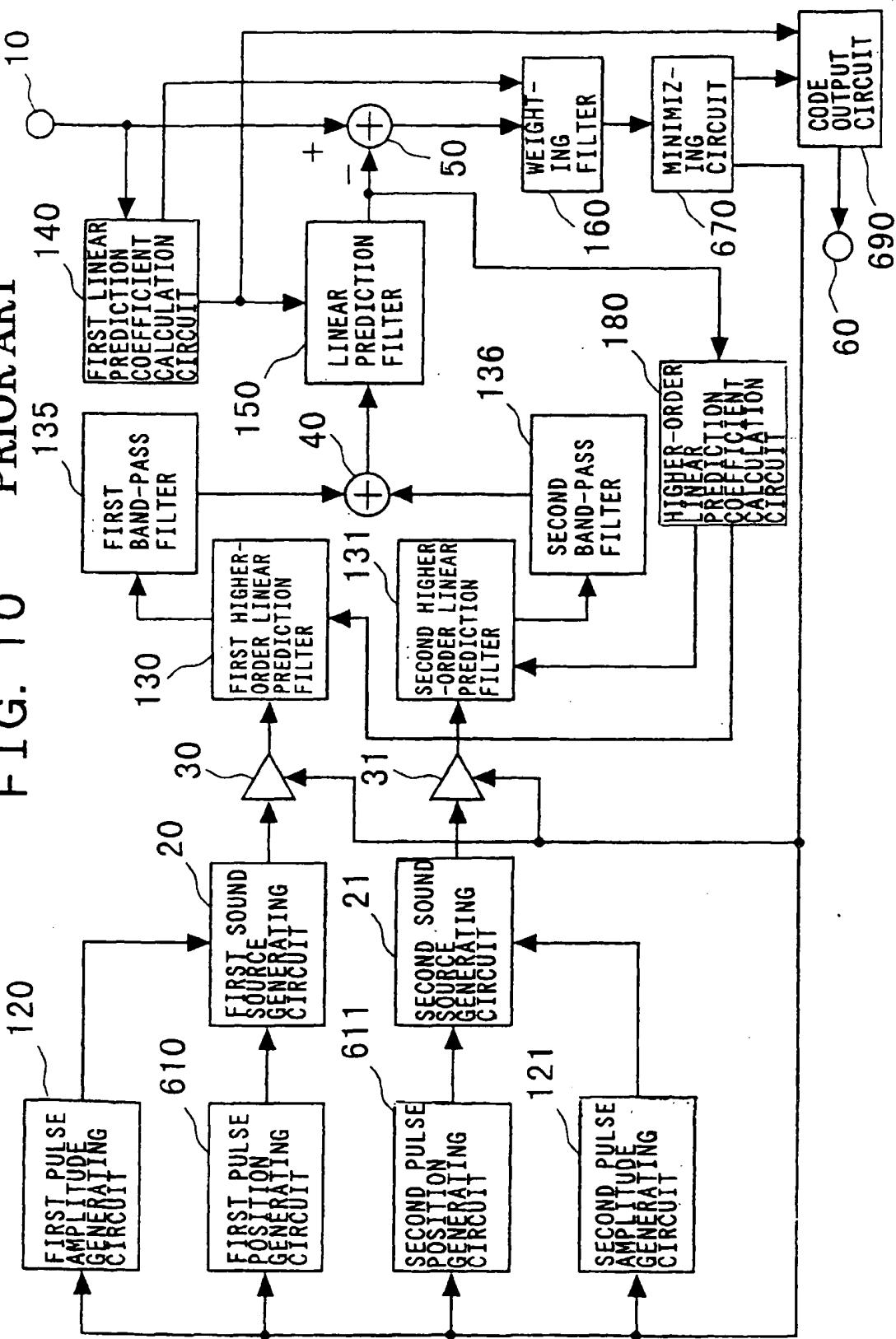
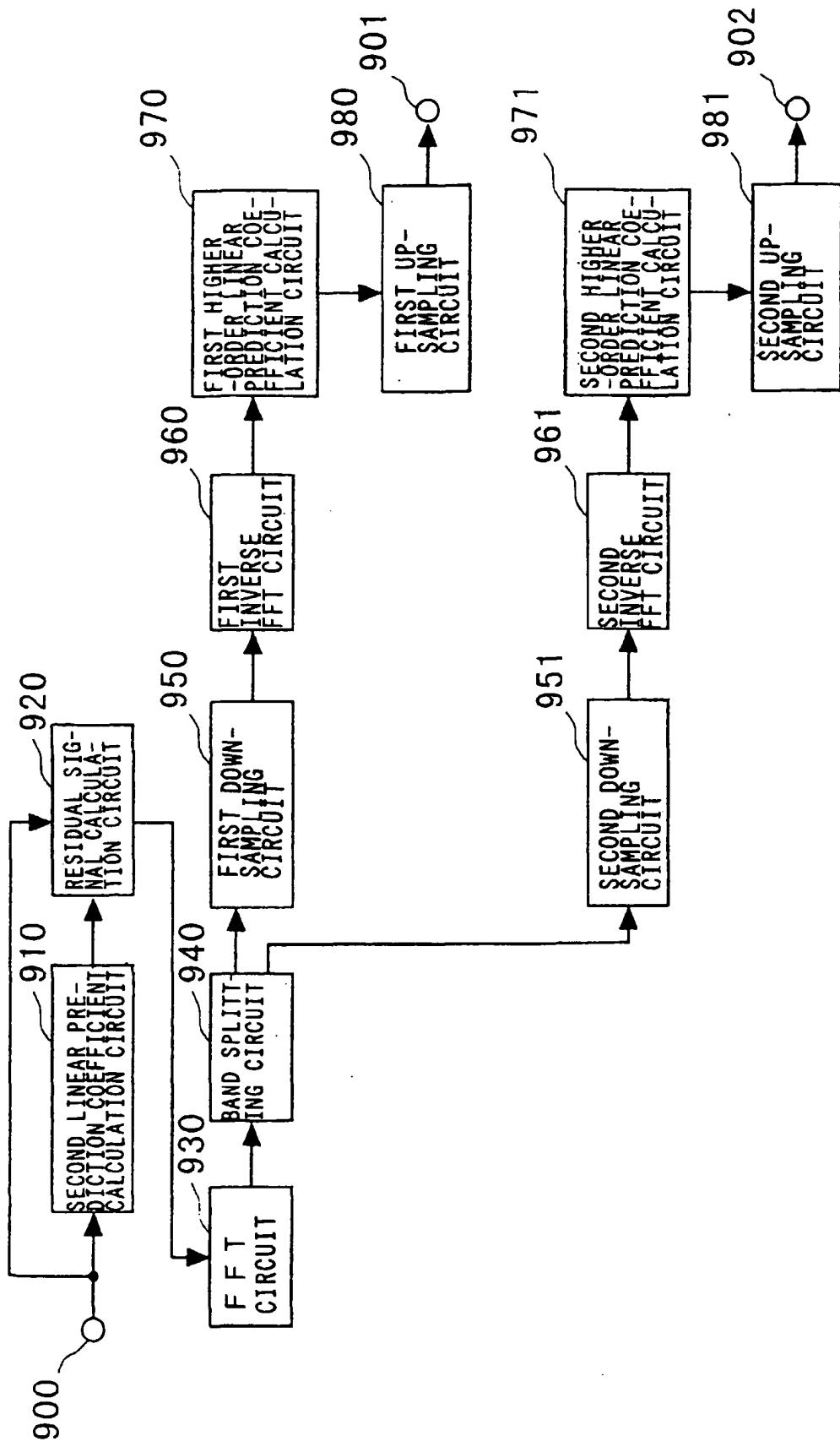
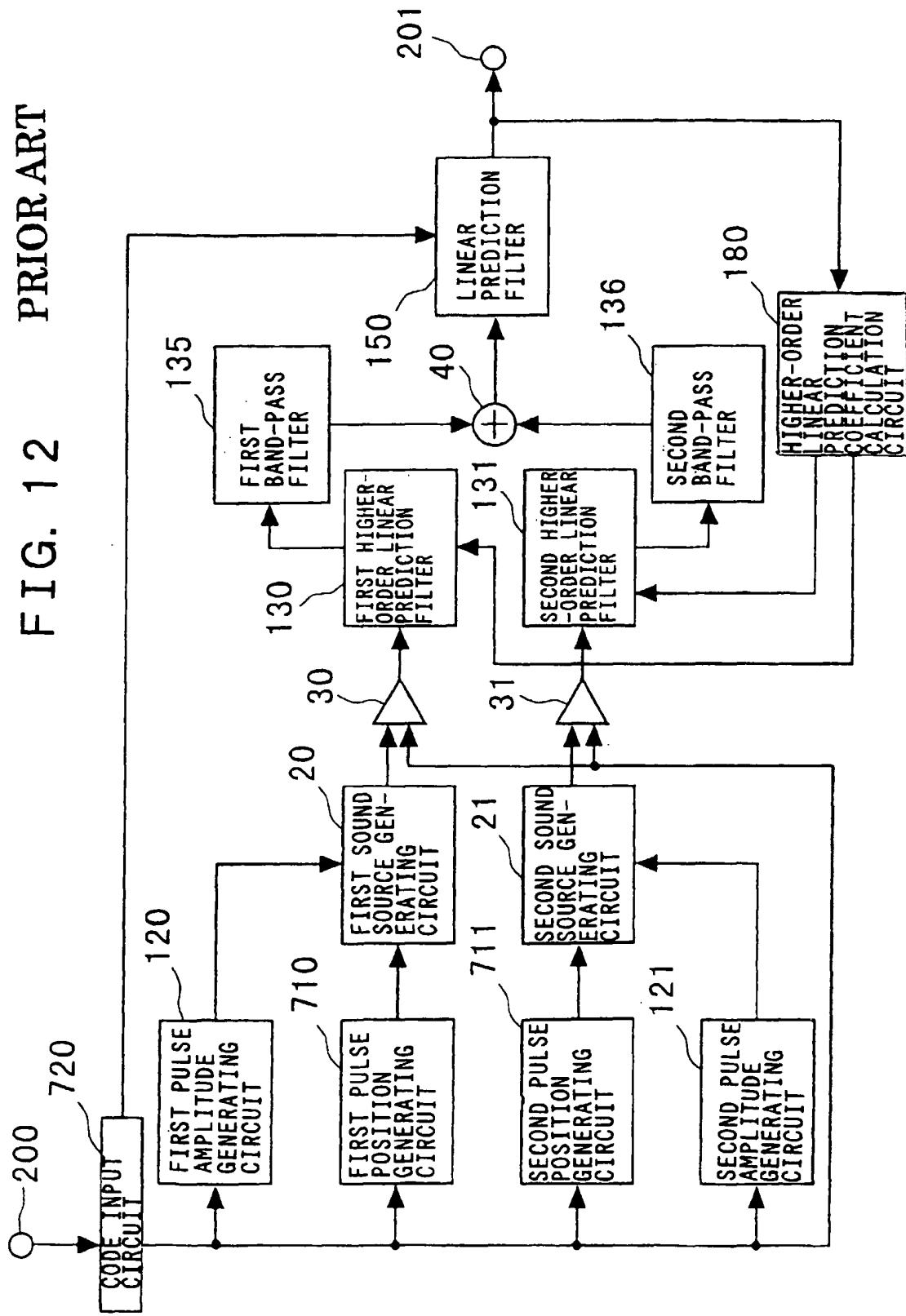


FIG. 11 PRIOR ART



## FIG. 12 PRIOR ART





DOCUMENTS CONSIDERED TO BE RELEVANT					
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)		
A	EP 0 396 121 A (CSELT CENTRO STUDI LAB TELECOM) 7 November 1990 * abstract; figure 1 * * page 2, line 37 - line 43 * ---	1-16	G10L9/14		
A	LIN X ET AL: "SUBBAND-MULTIPULSE DIGITAL AUDIO BROADCASTING FOR MOBILE RECEIVERS" IEEE TRANSACTIONS ON BROADCASTING, vol. 39, no. 4, 1 December 1993, pages 373-382, XP000514974 * Section 2 (pages 374-378) *	1-16			
A	OZAWA K ET AL: "MP-CELP SPEECH CODING BASED ON MULTIPULSE VECTOR QUANTIZATION AND FAST SEARCH" ELECTRONICS & COMMUNICATIONS IN JAPAN, PART III - FUNDAMENTAL ELECTRONIC SCIENCE, vol. 80, no. 11, 1 November 1997, pages 55-63, XP000754860 * Section 2 (pages 56-59) *	1-16			
D, A	UBALE A ET AL: "Multi-band CELP coding of speech and music" IEEE WORKSHOP ON SPEECH CODING FOR TELECOMMUNICATIONS. BACK TO BASICS: ATTACKING FUNDAMENTAL PROBLEMS IN SPEECH CODING, 1997, pages 101-102, XP002104207 IEEE, New York, NY, USA * the whole document *	1-16	<table border="1"> <tr> <td>TECHNICAL FIELDS SEARCHED (Int.Cl.6)</td> </tr> <tr> <td>G10L</td> </tr> </table>	TECHNICAL FIELDS SEARCHED (Int.Cl.6)	G10L
TECHNICAL FIELDS SEARCHED (Int.Cl.6)					
G10L					
<p>The present search report has been drawn up for all claims</p>					
Place of search  EPO FORM 1503 03/82 (P04001)	Date of completion of the search  28 May 1999	Examiner  Ramos Sánchez, U			
CATEGORY OF CITED DOCUMENTS		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document			
X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document					

ANNEX TO THE EUROPEAN SEARCH REPORT  
ON EUROPEAN PATENT APPLICATION NO.

EP 99 10 2792

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.  
The members are as contained in the European Patent Office EDP file on  
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

28-05-1999

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0396121 A	07-11-1990	CA 2016042 A,C DE 69028434 D DE 69028434 T	03-11-1990 17-10-1996 20-02-1997
-----			